

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
10 September 2004 (10.09.2004)

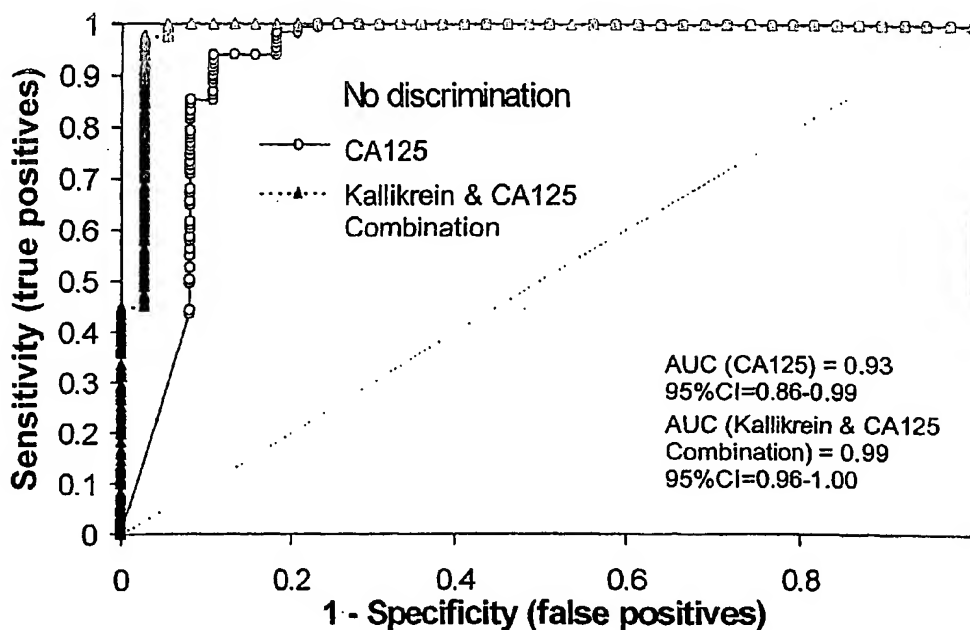
PCT

(10) International Publication Number  
**WO 2004/075713 A2**

- (51) International Patent Classification<sup>7</sup>: **A61B**
- (21) International Application Number:  
PCT/CA2004/000281
- (22) International Filing Date: 26 February 2004 (26.02.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/450,406 26 February 2003 (26.02.2003) US
- (71) Applicant (for all designated States except US): **MOUNT SINAI HOSPITAL** [CA/CA]; 600 University Avenue, Room 970, Toronto, Ontario M5G 1X5 (CA).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **DIAMANDIS, Eleftherios, P.** [CA/CA]; 44 Gerrard Street West, Suite 1504, Toronto, Ontario M5G 2K2 (CA).
- (74) Agent: **KURDYDYK, Linda**; McCarthy Tetrault, 66 Wellington Street West, Suite 4700, P.O. Box 48, Toronto Dominion Bank Tower (CA).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: MULTIPLE MARKER ASSAY FOR DETECTION OF OVARIAN CANCER



(57) Abstract: Methods for diagnosing and monitoring ovarian cancer in a subject comprising measuring a plurality of kallikrein polypeptides, and optionally CA125, or nucleic acids encoding the polypeptides in a sample from the subject. The kallikrein polypeptides include kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**TITLE: Multiple Marker Assay for Detection of Ovarian Cancer****FIELD OF THE INVENTION**

The invention relates to compositions, kits, and methods for detecting, characterizing, preventing, and treating ovarian cancer.

**5 BACKGROUND OF THE INVENTION**

Epithelial ovarian carcinoma is the most common and most lethal of all gynecologic malignancies. Only 30% of ovarian tumors are diagnosed at an early stage (Stage I/II), when survival rates reach 90%. The rest are diagnosed at an advanced stage, with survival rates of less than 20% ( Greenlee RT, Hill-Harmon MB, Murray T, et al., 2001. *CA Cancer J Clin* .2001;51:15-36). Currently, the only well-accepted  
 10 serological marker is CA125, a large glycoprotein of unknown function (Meyer T, Rustin GJ., *Br J Cancer* .2000;82:1535-1538). However, CA125 has limitations as a diagnostic, prognostic and screening tool (Holschneider CH, Berek JS, *Semin Surg Oncol* .2000;19:3-10). Consequently, there is a need to enhance the overall diagnostic/prognostic capability of CA125.

Kallikreins are a subgroup of secreted serine proteases, encoded by highly conserved and tightly  
 15 clustered multigene families in humans, rats and mice. The human kallikrein gene family resides on chromosome 19q13.4 and is comprised of 15 members, whose genes are designated as *KLK1* to *KLK15* and the corresponding proteins as hK1 to hK15 ( Yousef GM, Diamandis EP., *Endocr Rev* .2001;22:184-204; Yousef GM, Chang A, Scorilas A, et al., *Biochem Biophys Res Commun*. 2000;276:125-133; Diamandis EP, Yousef GM, Clements J, et al. *Clin Chem* .2000;46:1855-1858). Kallikreins are expressed in a wide variety  
 20 of tissues and are found in many biological fluids (e.g. cerebrospinal fluid, serum, seminal plasma, milk, etc.) where they are predicted to process specific substrates. Kallikreins may participate in cascade reactions similar to those involved in digestion, fibrinolysis, coagulation, wound healing and apoptosis (( Yousef GM, Diamandis EP., *Endocr Rev* .2001;22:184-204). Many kallikreins have been found to be differentially expressed in endocrine-related malignancies (Diamandis EP, Yousef GM, *Expert Rev. Mol. Diagn*  
 25 .2001;1:182-190), including prostate ( Barry MJ. Clinical practice, *N Engl J Med* .2001;344:1373-1377; Rittenhouse HG, Finlay JA, Mikolajczyk SD, et al., *Crit Rev Clin Lab Sci* .1998;35:275-368; and Yousef GM, Scorilas A, Jung K, et al., *J Biol Chem* .2001;276:53-61), ovarian ( Kim H, Scorilas A, Katsaros D, et al., *Br J Cancer*, 2001;84:643-650; Anisowicz A, Sotiropoulou G, Stenman, et al., *Mol Med* .1996;2:624-636; Tanimoto H, Underwood LJ, Shigemasa K, et al., *Cancer* .1999;86:2074-2082; Magklara A, Scorilas A, Katsaros D, et al., *Clin Cancer Res* .2001;7:806-811; Yousef GM, Kyriakopoulou LG, Scorilas A, et al.,  
 30 *Cancer Res* .2001;61:7811-7818; Luo L, Bunting P, Scorilas A, Diamandis EP., *Clin Chim Acta* .2001;306:111-118), breast ( Yousef GM, Magklara A, Chang A, et al., *Cancer Res* .2001;61:3425-3431; Yousef GM, Chang A, Diamandis EP; *J Biol Chem* .2000;275:11891-11898; and Yousef GM, Magklara A, Diamandis EP, *Genomics* .2000;69:331-341), and testicular cancer ( Luo LY, Rajpert-De Meyts ER, Jung K, et al.,2001;85:220-224). In addition, many kallikrein genes examined thus far are under steroid hormone regulation, implicating a role for kallikreins in endocrine-related tissues (Yousef GM, Diamandis EP., *Endocr Rev.*, 2001;22:184-204). Furthermore, hK6, hK10 and hK11 have been recently identified as novel serological ovarian cancer biomarkers ( Luo L, Bunting P, Scorilas A, Diamandis EP., *Clin Chim Acta* .2001;306:111-118 Diamandis EP, Yousef GM, Soosaipillai AR, Bunting P., *Clin Biochem*. 2000;33:579-

583, and Diamandis EP, Okui A, Mitsui S, et al., *Cancer Res.* 2002;62:295-300).

#### SUMMARY OF THE INVENTION

The present invention seeks to overcome the drawbacks inherent in the prior art and seeks to provide sensitive and accurate multimarker methods for the detection of ovarian cancer. A plurality of kallikrein polypeptides and polynucleotides encoding the polypeptides, optionally in combination with CA125 and polynucleotides encoding CA125 can have particular application in the detection of ovarian cancer. A plurality of kallikrein markers (i.e. two or more of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11) and polynucleotides encoding the polypeptides, optionally in combination with CA125 and polynucleotides encoding CA125, constitute biomarkers for the diagnosis, monitoring, progression, treatment, and prognosis of ovarian cancer, and they may be used as biomarkers before surgery or after relapse.

In accordance with the methods of the invention, the presence of levels of markers in a sample can be assessed, for example by detecting the presence in the sample of (a) polypeptides or polypeptide fragments corresponding to the markers; (b) metabolites which are produced directly or indirectly by polypeptides corresponding to the markers; (c) transcribed nucleic acids or fragments thereof having at least a portion with which the markers are substantially identical; and/or (c) transcribed nucleic acids or fragments thereof, wherein the nucleic acids hybridize with the markers.

In an aspect of the invention, a method is provided for detecting ovarian cancer in a patient comprising detecting a plurality of kallikrein polypeptides, optionally in combination with CA125, in a sample from the patient wherein the method provides substantially increased sensitivity compared to methods using CA125 alone. In an embodiment, sensitivity is increased by at least 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, and 35% compared to using CA125 alone.

In an embodiment, the invention provides a method for detecting a plurality of kallikrein markers, and optionally CA125, associated with ovarian cancer in a patient comprising:

- (a) obtaining a sample from a patient;
- (b) detecting or identifying in the sample kallikrein markers, optionally in combination with CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- (c) comparing the detected amounts with amounts detected for a standard.

The term "detect" or "detecting" includes assaying, assessing, imaging or otherwise establishing the presence or absence of the target kallikrein and CA125 polypeptides or polynucleotides encoding the polypeptides, subunits thereof, or combinations of reagent bound targets, and the like, or assaying for, imaging, ascertaining, establishing, or otherwise determining one or more factual characteristics of ovarian cancer, metastasis, stage, or similar conditions. The term encompasses diagnostic, prognostic, and monitoring applications. The kallikrein polypeptides and CA125 can be detected individually, sequentially, or simultaneously.

According to a method involving kallikrein markers optionally in combination with CA125, the levels in the sample of the kallikrein markers (2, 3, 4, 5, or 6) and optionally CA125, wherein the markers comprise or are selected from kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein



- 3 -

11, are compared with the normal levels of the kallikrein markers, and optionally CA125, in samples of the same type obtained from controls (e.g. samples from individuals not afflicted with ovarian cancer). Significantly different levels in the sample of the kallikrein markers (and optionally CA125) relative to the normal levels in a control is indicative of ovarian cancer.

5 In an embodiment, the invention provides a method for diagnosing and monitoring ovarian carcinoma in a subject comprising detecting in a sample from the subject kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The kallikrein markers and CA125 can be detected using antibodies that bind to the kallikrein markers and CA125 or parts thereof.

10 Thus, the invention provides a method of assessing whether a patient is afflicted with or has a predisposition for ovarian cancer, the method comprising comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a sample from the patient, wherein the kallikrein markers comprise kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 15 (b) normal levels of kallikrein markers, and optionally CA125, in samples of the same type obtained from control patients not afflicted with ovarian cancer, wherein significantly different levels of the kallikrein markers and optionally CA125, relative to the corresponding normal levels of the kallikrein markers, and optionally CA125, is an indication that the patient is afflicted with ovarian cancer.

20 In an embodiment of a method of assessing whether a patient is afflicted with ovarian cancer (e.g. screening, detection of a recurrence, reflex testing), the method comprises comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 25 (b) normal levels of the kallikrein markers, and optionally CA125, in a control non-ovarian cancer sample.

A significant difference between the levels of the kallikrein markers, and optionally CA125, in the patient sample and the normal levels is an indication that the patient is afflicted with ovarian cancer.

30 The invention further relates to a method of assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient. This method comprises comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a first sample obtained from the patient prior to providing at least a portion of the therapy to the patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11 ; and
- 35 (b) levels of the kallikrein markers, and optionally CA125, in a second sample obtained from the patient following therapy.

A significant difference between the levels of the kallikrein markers, and optionally CA125, in the second sample, relative to the first sample, is an indication that the therapy is efficacious for inhibiting ovarian cancer.

- 4 -

The "therapy" may be any therapy for treating ovarian cancer including but not limited to chemotherapy, immunotherapy, gene therapy, radiation therapy, and surgical removal of tissue. Therefore, the method can be used to evaluate a patient before, during, and after therapy, for example, to evaluate the reduction in tumor burden.

5 In an aspect, the invention provides a method for monitoring the progression of ovarian cancer in a patient, the method comprising:

- (a) detecting in a patient sample at a first time point, kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 10 (b) repeating step (a) at a subsequent point in time; and
- (c) comparing the levels detected in (a) and (b), and therefrom monitoring the progression of ovarian cancer in the patient.

In another aspect, the invention provides a method for assessing the aggressiveness or indolence of ovarian cancer (e.g. staging), the method comprising comparing:

- 15 (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- (b) normal levels of the kallikrein markers, and optionally CA125 in a control sample.

20 A significant difference between the levels in the sample and the normal levels is an indication that the cancer is aggressive or indolent.

The invention provides a method for determining whether ovarian cancer has metastasized or is likely to metastasize in the future, the method comprising comparing:

- 25 (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- (b) normal levels (or non-metastatic levels) of the kallikrein markers, and optionally CA125, in a control sample.

A significant difference between the levels in the patient sample and the normal levels is an indication that the cancer has metastasized or is likely to metastasize in the future.

30 The invention also provides a method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, and a method of selecting an agent for inhibiting ovarian cancer in a patient.

The invention further provides a method of inhibiting ovarian cancer in a patient comprising:

- 35 (a) obtaining a sample comprising cancer cells from the patient;
- (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents;
- (c) comparing levels of kallikrein markers, and optionally CA125, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11;
- (d) administering to the patient at least one of the test agents which alters the levels of the

- 5 -

kallikrein markers, and optionally CA125, in the aliquot containing that test agent, relative to other test agents.

The invention also contemplates a method of assessing the ovarian carcinogenic potential of a test compound comprising:

- 5 (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and
- (b) comparing levels of kallikrein markers, and optionally CA125, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

10 A significant difference between the levels of the kallikrein markers, and optionally CA125, in the aliquot maintained in the presence of (or exposed to) the test compound relative to the aliquot maintained in the absence of the test compound, indicates that the test compound possesses ovarian carcinogenic potential.

In preferred embodiments of the methods of the invention, the kallikrein markers comprise a plurality of kallikrein markers, for example, at least three, four, five, or six of the markers. In particular, a  
15 plurality of kallikrein markers may be selected from the group consisting of kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10, from the group consisting of kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, or from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

Other methods of the invention employ one or more polynucleotides capable of hybridizing to  
20 polynucleotides encoding kallikrein markers, and optionally CA125. Methods for detecting polynucleotides encoding a kallikrein markers, and optionally CA125, can be used to monitor ovarian cancer by detecting the nucleic acids.

Thus, the present invention relates to a method for diagnosing and monitoring ovarian cancer in a sample from a subject comprising isolating nucleic acids, preferably mRNA, from the sample; and detecting  
25 polynucleotides encoding kallikrein markers, and optionally CA125, in the sample. The presence of different levels of polynucleotides encoding kallikrein markers, and optionally CA125, in the sample compared to a standard or control is indicative of disease, disease stage, and/or prognosis, e.g. longer progression-free and overall survival.

In an embodiment, the invention provides methods for determining the presence or absence of  
30 ovarian cancer in a subject comprising (a) contacting a sample obtained from the subject with oligonucleotides that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125; and (b) detecting in the sample levels of nucleic acids that hybridize to the polynucleotides relative to a predetermined cut-off value, and therefrom determining the presence or absence of ovarian cancer in the subject. Within certain embodiments, mRNA is detected via polymerase chain reaction using, for example  
35 oligonucleotide primers that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or complements of such polynucleotides. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing oligonucleotide probes that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or complements of such polynucleotides.

When using mRNA detection, the method may be carried out by combining isolated mRNA with

- 6 -

reagents to convert to cDNA according to standard methods; treating the converted cDNA with amplification reaction reagents (such as cDNA PCR reaction reagents) in a container along with an appropriate mixture of nucleic acid primers; reacting the contents of the container to produce amplification products; and analyzing the amplification products to detect the presence of polynucleotides encoding kallikrein markers, and optionally CA125, in the sample. For mRNA the analyzing step may be accomplished using Northern Blot analysis to detect the presence of polynucleotides encoding kallikrein markers, and optionally CA125. The analysis step may be further accomplished by quantitatively detecting the presence of polynucleotides encoding kallikrein markers, and optionally CA125, in the amplification product, and comparing the quantity of markers detected against a panel of expected values for the known presence or absence of the kallikrein markers in normal and malignant tissue derived using similar primers.

In embodiments of the methods of the invention, a plurality (eg. three, four, five or six) polynucleotides encoding kallikrein polypeptides are employed. In particular, a plurality of polynucleotides encoding kallikrein markers may be selected from the group consisting of polynucleotides encoding (i) kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; (ii) polynucleotides encoding kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (iii) polynucleotides encoding kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

The invention also provides a diagnostic composition comprising a plurality of kallikrein polypeptides and optionally CA125 polypeptide, or polynucleotides encoding the polypeptides, or agents that bind to the polypeptides or polynucleotides.

In an embodiment, the composition comprises probes that specifically hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or fragments thereof. In another embodiment a composition is provided comprising specific primer pairs capable of amplifying polynucleotides encoding kallikrein markers, and optionally CA125, using polymerase chain reaction methodologies. In a still further embodiment, the composition comprises agents that bind to kallikrein markers, and optionally CA125, (e.g. antibodies) or fragments thereof. Probes, primers, and agents can be labeled with detectable substances.

In an aspect the invention provides an *in vivo* method comprising administering to a subject agents that have been constructed to target kallikrein markers, and optionally CA125.

The invention therefore contemplates an *in vivo* method comprising administering to a mammal imaging agents that carry labels for imaging and that bind to kallikrein markers, and optionally CA125, and then imaging the mammal.

Still further the invention relates to therapeutic applications for ovarian cancer employing kallikrein markers, and optionally CA125, nucleic acids encoding the polypeptides, and/or agents identified using methods of the invention.

The invention also includes kits for carrying out methods of the invention. In an embodiment, the kit is for assessing whether a patient is afflicted with ovarian cancer and it comprises reagents for assessing kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

In another aspect the invention relates to a kit for assessing the suitability of each of a plurality of test compounds for inhibiting ovarian cancer in a patient. The kit comprises reagents for assessing kallikrein

- 7 -

markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The kit may also comprise a plurality of test agents or compounds.

5 The invention contemplates a kit for assessing the presence of ovarian cancer cells, wherein the kit comprises antibodies specific for selected kallikrein markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

10 Additionally the invention provides a kit for assessing the ovarian carcinogenic potential of a test compound. The kit comprises ovarian cells and reagents for assessing kallikrein markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

15 In an aspect the invention provides a method of treating a patient afflicted with ovarian cancer comprising providing to cells of a patient antisense oligonucleotides complementary to polynucleotides encoding kallikrein markers, and optionally CA125, which are overexpressed in ovarian cancer. In an alternative method, expression of genes corresponding to kallikrein markers, and optionally CA125, which are underexpressed in ovarian cancer are increased.

20 The invention relates to a method of inhibiting ovarian cancer in a patient at risk for developing ovarian cancer comprising inhibiting or increasing expression (or overexpression) of genes encoding kallikrein markers and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, that are either overexpressed or underexpressed, in ovarian cancer.

25 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

#### **DESCRIPTION OF THE DRAWINGS**

The invention will now be described in relation to the drawings in which

30 Figure 1 is a graph showing hk5 concentration in serum from non-cancer and cancer patients.  
Figure 2 is a graph showing hk6 concentration in serum from non-cancer and cancer patients.  
Figure 3 is a graph showing hk7 concentration in serum from non-cancer and cancer patients.  
Figure 4 is a graph showing hk8 concentration in serum from non-cancer and cancer patients.  
Figure 5 is a graph showing hk10 concentration in serum from non-cancer and cancer patients.  
Figure 6 is a graph showing hk11 concentration in serum from non-cancer and cancer patients.  
35 Figure 7 is a graph showing CA125 concentration in serum from non-cancer and cancer patients.  
Figure 8 is a ROC curve illustrating the added value of using kallikreins and CA125 together in a multivariate function.

#### **DETAILED DESCRIPTION OF THE INVENTION**

The invention relates to newly discovered correlations between expression of certain markers and

ovarian cancer. The combinations of markers described herein may provide sensitive methods for detecting ovarian cancer. The levels of expression of a combination of markers described herein may correlate with the presence of ovarian cancer or a pre-malignant condition in a patient. Methods are provided for detecting the presence of ovarian cancer in a sample, the absence of ovarian cancer in a sample, the stage of an ovarian cancer, the grade of an ovarian cancer, the benign or malignant nature of an ovarian cancer, the metastatic potential of an ovarian cancer, assessing the histological type of neoplasm associated with the ovarian cancer, the indolence or aggressiveness of the cancer, and other characteristics of ovarian cancer that are relevant to prevention, diagnosis, characterization, and therapy of ovarian cancer in a patient. Methods are also provided for assessing the efficacy of one or more test agents for inhibiting ovarian cancer, assessing the efficacy of a therapy for ovarian cancer, monitoring the progression of ovarian cancer, selecting an agent or therapy for inhibiting ovarian cancer, treating a patient afflicted with ovarian cancer, inhibiting ovarian cancer in a patient, and assessing the carcinogenic potential of a test compound.

#### Glossary

The terms "sample", "biological sample", and the like, mean a material known or suspected of expressing or containing a plurality of kallikrein markers or polypeptides (2, 3, 4, 5, or 6 polypeptides), and optionally CA125 polypeptide, or polynucleotides encoding the polypeptides. The test sample can be used directly as obtained from the source or following a pretreatment to modify the character of the sample. The sample can be derived from any biological source, such as tissues, extracts, or cell cultures, including cells (e.g. tumor cells), cell lysates, and physiological fluids, such as, for example, whole blood, plasma, serum, saliva, ocular lens fluid, cerebral spinal fluid, sweat, urine, milk, ascites fluid, synovial fluid, peritoneal fluid and the like. The sample can be obtained from animals, preferably mammals, most preferably humans. The sample can be treated prior to use, such as preparing plasma from blood, diluting viscous fluids, and the like. Methods of treatment can involve filtration, distillation, extraction, concentration, inactivation of interfering components, the addition of reagents, and the like. Nucleic acids and polypeptides may be isolated from the samples and utilized in the methods of the invention. In a preferred embodiment, the sample is a serum sample.

The term "subject" or "patient" refers to a warm-blooded animal such as a mammal, which is suspected of having ovarian cancer, or a condition, disease, or syndrome associated with ovarian cancer. Preferably, "subject" refers to a human.

"CA125", "CA125 polypeptide", or "carbohydrate antigen 125" refers to a high-molecular weight mucin, which can be defined by its ability to bind to monoclonal antibody OC125. The CA125 protein core comprises a short cytoplasmic core tail, a transmembrane domain, and a large and heavily glycosylated extracellular domain dominated by a repeat domain of 156 amino acids rich in serine, threonine, and proline (Yin BW and Lloyd KO, J Biol Chem. 2001, 276:27371-27375; O'Brian TJ et al, Tumor Biol., 2001 22:348-366; and Hovig E. et al, Tumor Biol. 2001, 22:345-347). The sequence of CA125 is shown in GenBank Accession No. NP\_078966, AAL65133 and AF414442 (SEQ ID NO. 1). The term includes the native-sequence polypeptides, isoforms, precursors and chimeric polypeptides. The term also includes the native sequence polypeptide, including polypeptide variants and polypeptides with substantial sequence identity (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or

99% sequence identity) to the sequence of GenBank Accession No.NP\_078966 (SEQ ID NO. 1), and that preferably retain the immunogenic activity of the corresponding native sequence polypeptide.

“Kallikrein polypeptides” or “kallikrein markers” comprise kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The term includes the native-sequence polypeptides, isoforms, precursors and chimeric polypeptides. The amino acid sequences for native kallikrein polypeptides employed in the present invention include the sequences found in GenBank for each polypeptide as shown in Table 1, and in SEQ ID NO: 3 (kallikrein 5), NO.6 (kallikrein 6), NO. 10 (kallikrein 7), NO. 13 (kallikrein 8), NO. 16 (kallikrein 10), and NOs. 19 and 20 (kallikrein 11), or a portion thereof. Other useful polypeptides are substantially identical to these sequences (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity), and preferably retain the immunogenic activity of the corresponding native-sequence kallikrein polypeptide.

A “native-sequence polypeptide” comprises a polypeptide having the same amino acid sequence of a polypeptide derived from nature. Such native-sequence polypeptides can be isolated from nature or can be produced by recombinant or synthetic means.

The term “native-sequence polypeptide” specifically encompasses naturally occurring truncated or secreted forms of a polypeptide, polypeptide variants including naturally occurring variant forms (e.g., alternatively spliced forms or splice variants), and naturally occurring allelic variants.

The term “polypeptide variant” means a polypeptide having at least about 70-80%, preferably at least about 85%, more preferably at least about 90%, most preferably at least about 95% amino acid sequence identity with a native-sequence polypeptide, in particular having at least 70-80%, 85%, 90%, 95% amino acid sequence identity to the sequences identified in the GenBank Accession Nos. in Table 1 and Accession No. NP\_078966, AF414442 and AAL65133 and shown in SEQ ID NOS: 1, 2, 3, 6, 10, 13, 16, 19 and 20. Such variants include, for instance, polypeptides wherein one or more amino acid residues are added to, or deleted from, the N- or C-terminus of the full-length or mature sequences of SEQ ID NOS: 1, 2, 3, 6, 10, 13, 16, 19 and 20, including variants from other species, but excludes a native-sequence polypeptide.

An allelic variant may also be created by introducing substitutions, additions, or deletions into a nucleic acid encoding a native polypeptide sequence such that one or more amino acid substitutions, additions, or deletions are introduced into the encoded protein. Mutations may be introduced by standard methods, such as site-directed mutagenesis and PCR-mediated mutagenesis. In an embodiment, conservative substitutions are made at one or more predicted non-essential amino acid residues. A “conservative amino acid substitution” is one in which an amino acid residue is replaced with an amino acid residue with a similar side chain. Amino acids with similar side chains are known in the art and include amino acids with basic side chains (e.g. Lys, Arg, His), acidic side chains (e.g. Asp, Glu), uncharged polar side chains (e.g. Gly, Asp, Glu, Ser, Thr, Tyr and Cys), nonpolar side chains (e.g. Ala, Val, Leu, Iso, Pro, Trp), beta-branched side chains (e.g. Thr, Val, Iso), and aromatic side chains (e.g. Tyr, Phe, Trp, His). Mutations can also be introduced randomly along part or all of the native sequence, for example, by saturation mutagenesis. Following mutagenesis the variant polypeptide can be recombinantly expressed and the activity of the polypeptide may be determined.

Polypeptide variants include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of a native polypeptide which include fewer amino acids than the full length polypeptides. A portion of a polypeptide can be a polypeptide which is for example, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more amino acids in length. Portions in which regions of a polypeptide are deleted can be prepared by recombinant techniques and can be evaluated for one or more functional activities such as the ability to form antibodies specific for a polypeptide.

A naturally occurring allelic variant may contain conservative amino acid substitutions from the native polypeptide sequence or it may contain a substitution of an amino acid from a corresponding position in a CA125 or kallikrein polypeptide homolog, for example, the murine CA125 or kallikrein polypeptide.

Percent identity of two amino acid sequences, or of two nucleic acid sequences identified herein is defined as the percentage of amino acid residues or nucleotides in a candidate sequence that are identical with the amino acid residues in a CA125 or kallikrein polypeptide or nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid or nucleic acid sequence identity can be achieved in various conventional ways, for instance, using publicly available computer software including the GCG program package (Devereux J. et al., *Nucleic Acids Research* 12(1): 387, 1984); BLASTP, BLASTN, and FASTA (Altschul, S.F. et al. *J. Molec. Biol.* 215: 403-410, 1990). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S. et al. NCBI NLM NIH Bethesda, Md. 20894; Altschul, S. et al. *J. Mol. Biol.* 215: 403-410, 1990). Skilled artisans can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Methods to determine identity and similarity are codified in publicly available computer programs.

CA125 and kallikrein polypeptides include chimeric or fusion proteins. A "chimeric protein" or "fusion protein" comprises all or part (preferably biologically active) of a CA125 or kallikrein polypeptide operably linked to a heterologous polypeptide (i.e., a polypeptide other than the same CA125 or kallikrein polypeptide). Within the fusion protein, the term "operably linked" is intended to indicate that the CA125 or kallikrein polypeptide and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the N-terminus or C-terminus of the CA125 or kallikrein polypeptide. A useful fusion protein is a GST fusion protein in which a kallikrein polypeptide is fused to the C-terminus of GST sequences. Another example of a fusion protein is an immunoglobulin fusion protein in which all or part of a CA125 or kallikrein polypeptide is fused to sequences derived from a member of the immunoglobulin protein family. Chimeric and fusion proteins can be produced by standard recombinant DNA techniques.

CA125 and kallikrein polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods, or by any combination of these and similar techniques.

"CA125 polynucleotides" or "polynucleotides encoding CA125" include nucleic acids that encode a native-sequence polypeptide, a polypeptide variant including a portion of a CA125 polypeptide, an isoform, precursor, and chimeric polypeptide. A nucleic acid sequence encoding native CA125 employed in the



- 11 -

present invention includes the nucleic acid sequence in GenBank Accession No. AF414442 and SEQ ID NO. 2, or a fragment thereof.

“Kallikrein polynucleotides” or “polynucleotides encoding kallikrein markers/polypeptides” refers to kallikrein 5 nucleic acids (KLK5), kallikrein 6 nucleic acids (KLK6), kallikrein 7 nucleic acids (KLK7),  
5 kallikrein 8 nucleic acids (KLK8), kallikrein 10 nucleic acids (KLK10), and/or kallikrein 11 nucleic acids (KLK11). The term includes nucleic acids that encode a native-sequence polypeptide, a polypeptide variant including a portion of a kallikrein polypeptide, an isoform, precursor, and chimeric polypeptide.

The polynucleotide sequences encoding native kallikrein polypeptides employed in the present invention include the nucleic acid sequences of the GenBank Accession Nos. identified in Table 1, and in  
10 SEQ ID NOS: 4 and 5 (KLK5), NOS. 7, 8, and 9 (KLK6), NOS. 11 and 12 (KLK 7), NOS. 14 and 15 (KLK8), NOS. 17 and 18 (KLK10), and NOS. 21 and 22 (KLK11), or a fragment thereof.

Polynucleotides encoding kallikrein polypeptides and CA125 include nucleic acid sequences complementary to these polynucleotides, and polynucleotides that are substantially identical to these sequences (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%,  
15 98%, or 99% sequence identity).

CA125 and kallikrein polynucleotides also include sequences which differ from a nucleic acid sequence of GenBank Accession Nos. identified in Table 1 and SEQ ID NOS: 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 17, 18, 21, and 22, due to degeneracy in the genetic code. As one example, DNA sequence polymorphisms within the nucleotide sequence of a CA125 or kallikrein polypeptide may result in silent mutations which do  
20 not affect the amino acid sequence. Variations in one or more nucleotides may exist among individuals within a population due to natural allelic variation. DNA sequence polymorphisms may also occur which lead to changes in the amino acid sequence of CA125 or a kallikrein polypeptide.

CA125 and kallikrein polynucleotides also include nucleic acids that hybridize under stringent conditions, preferably high stringency conditions to a nucleic acid sequence of the GenBank Accession Nos. identified in Table 1 and SEQ ID NOS: 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 17, 18, 21, and 22. Appropriate stringency conditions which promote DNA hybridization are known to those skilled in the art, or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C may be employed. The stringency may be selected based on the conditions used in the wash step. By  
30 way of example, the salt concentration in the wash step can be selected from a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be at high stringency conditions, at about 65°C.

CA125 and kallikrein polynucleotides also include truncated nucleic acids or fragments and variant forms of the polynucleotides that arise by alternative splicing of an mRNA corresponding to a DNA.

35 The CA125 and kallikrein polynucleotides are intended to include DNA and RNA (e.g. mRNA) and can be either double stranded or single stranded. A polynucleotide may, but need not, include additional coding or non-coding sequences, or it may, but need not, be linked to other molecules and/or carrier or support materials. The polynucleotides for use in the methods of the invention may be of any length suitable for a particular method.

A plurality of kallikrein polypeptides or kallikrein polynucleotides are generally detected in the present invention. "Plurality" refers to 2, 3, 4, 5, or 6 kallikrein polypeptides or polynucleotides, in particular 3, 4, 5, or 6, preferably 4, 5, or 6, more preferably 5 or 6 kallikrein polypeptides or polynucleotides.

In an embodiment a plurality of kallikrein polypeptides is selected from the group consisting of  
5 kallikrein 5, kallikrein 7, and kallikrein 8; kallikrein 5, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; or kallikrein 5, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. In another embodiment, a plurality of kallikrein polypeptides is selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

10 In an embodiment, a plurality of kallikrein polynucleotides is selected from the group consisting of KLK5, KLK7, and KLK8; KLK5, KLK8 and KLK10; KLK7, KLK8 and KLK10; KLK5, KLK7, KLK8, and KLK10; KLK7, KLK8, KLK10 and KLK11, or KLK5, KLK7, KLK8, KLK10 and KLK11. In another embodiment, a plurality of kallikrein polynucleotides is selected from the group consisting of KLK5, KLK6, KLK7, KLK8, KLK10, and KLK11.

#### 15 **General Methods**

A variety of methods can be employed for the diagnostic and prognostic evaluation of ovarian cancer involving kallikrein polypeptides, and optionally CA125 polypeptide, and polynucleotides encoding the polypeptides, and the identification of subjects with a predisposition to such disorders. Such methods may, for example, utilize polynucleotides encoding kallikrein polypeptides, and optionally CA125, and  
20 fragments thereof, and binding agents (e.g. antibodies aptamers) against kallikrein polypeptides, and optionally CA125 polypeptide, including peptide fragments. In particular, the polynucleotides and antibodies may be used, for example, for (1) the detection of either over- or under-expression of kallikrein polynucleotides, and optionally CA125, relative to a non-disorder state; and (2) the detection of either an over- or an under-abundance of kallikrein polypeptides, and optionally CA125, relative to a non-disorder  
25 state or the presence of modified (e.g., less than full length) kallikrein polypeptides, and optionally CA125, that correlate with a disorder state, or a progression toward a disorder state.

The invention also contemplates a method for detecting ovarian cancer comprising producing a profile of levels of a plurality of kallikrein markers, and optionally CA125, in cells from a patient, wherein the markers are kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and  
30 comparing the profile with a reference to identify a protein profile for the test cells indicative of disease.

The methods described herein may be used to evaluate the probability of the presence of malignant or pre-malignant cells, for example, in a group of cells freshly removed from a host. Such methods can be used to detect tumors, quantitate their growth, and help in the diagnosis and prognosis of disease. The methods can be used to detect the presence of cancer metastasis, as well as confirm the absence or removal  
35 of all tumor tissue following surgery, cancer chemotherapy, and/or radiation therapy. They can further be used to monitor cancer chemotherapy and tumor reappearance.

The methods described herein can be adapted for diagnosing and monitoring ovarian cancer by detecting a plurality of kallikrein polypeptides, and optionally CA125 polypeptide, or nucleic acids encoding the polypeptides in biological samples from a subject. These applications require that the amount of

- 13 -

polypeptides or nucleic acids quantitated in a sample from a subject being tested be compared to a predetermined standard. The standard may correspond to levels quantitated for another sample or an earlier sample from the subject, or levels quantitated for a control sample. Levels for control samples from healthy subjects or ovarian cancer subjects may be established by prospective and/or retrospective statistical studies.

- 5 Healthy or normal subjects who have no clinically evident disease or abnormalities may be selected for statistical studies. Diagnosis may be made by a finding of statistically different levels of a plurality of kallikrein polypeptides, and optionally CA125, or nucleic acids encoding same, compared to a control sample or previous levels quantitated for the same subject. A "significant difference" in levels of kallikrein markers or polynucleotides encoding the kallikrein markers in a patient sample compared to a control or  
10 standard (e.g. normal levels or levels in other samples from a patient) may represent levels that are higher or lower than the standard error of the detection assay, preferably the levels are at least about 1.5, 2, 3, 4, 5, or 6 times higher or lower, respectively, than the control or standard. The difference in levels of markers or polynucleotides may be a "statistically significant difference"

#### Nucleic Acid Methods/Assays

- 15 As noted herein an ovarian cancer may be detected based on the levels of polynucleotides encoding kallikrein polypeptides, and optionally CA125, in a sample. Techniques for detecting polynucleotides such as polymerase chain reaction (PCR) and hybridization assays are well known in the art.

- Nucleotide probes for use in the detection of nucleic acid sequences in samples may be constructed using conventional methods known in the art. Suitable probes may be based on nucleic acid sequences  
20 encoding at least 5 sequential amino acids from regions of nucleic acids encoding kallikrein polypeptides, and optionally CA125, preferably they comprise 15 to 40 nucleotides. A nucleotide probe may be labeled with a detectable substance such as a radioactive label that provides for an adequate signal and has sufficient half-life such as  $^{32}\text{P}$ ,  $^3\text{H}$ ,  $^{14}\text{C}$  or the like. Other detectable substances that may be used include antigens that are recognized by a specific labeled antibody, fluorescent compounds, enzymes, antibodies specific for a  
25 labeled antigen, and luminescent compounds. An appropriate label may be selected having regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. Labeled probes may be hybridized to nucleic acids on solid supports such as nitrocellulose filters or nylon membranes as generally described in Sambrook et al, 1989, Molecular Cloning, A Laboratory Manual (2nd ed.). The nucleic acid probes may be used to detect polynucleotides  
30 encoding kallikrein polypeptides, and optionally CA125, preferably in human cells. The nucleotide probes may also be useful in the diagnosis of ovarian cancer involving polynucleotides encoding kallikrein polypeptides, and optionally CA125, in monitoring the progression of such disorder; or monitoring a therapeutic treatment.

- Probes may be used in hybridization techniques to detect nucleic acids encoding a plurality of  
35 kallikrein polypeptides, and optionally CA125. The technique generally involves contacting and incubating nucleic acids (e.g. recombinant DNA molecules, cloned genes) obtained from a sample from a patient or other cellular source with probes under conditions favorable for the specific annealing of the probes to complementary sequences in the nucleic acids. After incubation, the non-annealed nucleic acids are removed, and the presence of nucleic acids that have hybridized to the probe if any are detected.

- 14 -

The detection of polynucleotides encoding kallikrein polypeptides and optionally CA125, may involve the amplification of specific gene sequences using an amplification method such as polymerase chain reaction (PCR), followed by the analysis of the amplified molecules using techniques known to those skilled in the art. Suitable primers can be routinely designed by one of skill in the art.

5 By way of example, oligonucleotide primers may be employed in a PCR based assay to amplify a portion of nucleic acids encoding each of a plurality of kallikrein polypeptides, and optionally CA125, derived from a sample, wherein the oligonucleotide primers are specific for (i.e. hybridize to) polynucleotides encoding each of the plurality of kallikrein polypeptides, and optionally CA125. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel  
10 electrophoresis.

In order to maximize hybridization under assay conditions, primers and probes employed in the methods of the invention generally have at least about 60%, preferably at least about 75% and more preferably at least about 90% identity to a portion of polynucleotides encoding a plurality of kallikrein polypeptides, and CA125. The primers and probes may be at least 10 nucleotides, and preferably at least 20  
15 nucleotides in length. In an embodiment the primers and probes are at least about 10-40 nucleotides in length.

Hybridization and amplification techniques described herein may be used to assay qualitative and quantitative aspects of expression of polynucleotides encoding kallikrein polypeptides, and optionally CA125. For example, RNA may be isolated from a cell type or tissue known to express these  
20 polynucleotides and tested utilizing the hybridization (e.g. standard Northern analyses) or PCR techniques referred to herein.

The primers and probes may be used in the above-described methods *in situ* i.e. directly on tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections.

In an aspect of the invention, a method is provided employing reverse transcriptase-polymerase  
25 chain reaction (RT-PCR), in which PCR is applied in combination with reverse transcription. Generally, RNA is extracted from a sample tissue using standard techniques (for example, guanidine isothiocyanate extraction as described by Chomcynski and Sacchi, Anal. Biochem. 162:156-159, 1987) and is reverse transcribed to produce cDNA. The cDNA is used as a template for a polymerase chain reaction. The cDNA is hybridized to sets of primers specifically designed against each of a plurality of kallikrein polynucleotide  
30 sequences, and optionally CA125. Once the primer and template have annealed a DNA polymerase is employed to extend from the primer, to synthesize a copy of the template. The DNA strands are denatured, and the procedure is repeated many times until sufficient DNA is generated to allow visualization by ethidium bromide staining and agarose gel electrophoresis.

Amplification may be performed on samples obtained from a subject with suspected ovarian cancer  
35 and an individual who is not afflicted with ovarian cancer. The reaction may be performed on several dilutions of cDNA spanning at least two orders of magnitude. A statistically significant difference in expression in several dilutions of the subject sample as compared to the same dilutions of the non-cancerous sample may be considered positive for the presence of ovarian cancer.

Oligonucleotides or longer fragments derived from polynucleotides encoding each of a plurality of

kallikrein polypeptides and optionally CA125, may be used as targets in a microarray. The microarray can be used to simultaneously monitor the expression levels of large numbers of genes. The information from the microarray may be used to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

5       The preparation, use, and analysis of microarrays are well known to a person skilled in the art. (See, for example, Brennan, T. M. et al. (1995) U.S. Pat. No. 5,474,796; Schena, et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995), PCT Application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R. A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M. J. et al. (1997) U.S. Pat. No. 5,605,662.)

10       Thus, the invention also includes an array comprising a plurality of polynucleotides encoding kallikrein marker(s), and optionally CA125 polynucleotides. The array can be used to assay expression of kallikrein polynucleotides, and optionally CA125 polynucleotides in the array. The invention allows the quantitation of expression of a plurality of kallikrein polynucleotides, and optionally CA125 polynucleotides.

15       In an embodiment, the array can be used to monitor the time course of expression of a plurality of kallikrein polynucleotides, and optionally CA125 polynucleotides, in the array. This can occur in various biological contexts such as tumor progression.

      The array is also useful for ascertaining differential expression patterns of a plurality of kallikrein polynucleotides and optionally CA125 polynucleotides, in normal and abnormal cells. This provides a battery of polynucleotides that could serve as molecular targets for diagnosis or therapeutic intervention.

20       **Protein Methods**

      Binding agents specific for a plurality of kallikrein markers and CA125 may be used for a variety of diagnostic and assay applications. There are a variety of assay formats known to the skilled artisan for using a binding agent to detect a target molecule in a sample. (For example, see Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988). In general, the presence or absence of an ovarian cancer in a subject may be determined by (a) contacting a sample from the subject with binding agents for a plurality of kallikrein polypeptides, and optionally CA125; (b) detecting in the sample levels of polypeptides that bind to the binding agents; and (c) comparing the levels of polypeptides with a predetermined standard or cut-off value.

30       "Binding agent" refers to a substance such as a polypeptide or antibody that specifically binds to a kallikrein or CA125 polypeptide. A substance "specifically binds" to a polypeptide if it reacts at a detectable level with the kallikrein or CA125 polypeptide, and does not react detectably with peptides containing unrelated sequences or sequences of different polypeptides. Binding properties may be assessed using an ELISA, which may be readily performed by those skilled in the art (see for example, Newton et al, Develop. Dynamics 197: 1-13, 1993).

35       A binding agent may be a ribosome, with or without a peptide component, an aptamer, an RNA molecule, or a polypeptide. A binding agent may be a polypeptide that comprises a kallikrein polypeptide or CA125 polypeptide sequence, a peptide variant thereof, or a non-peptide mimetic of such a sequence. By way of example a kallikrein polypeptide sequence may be a peptide portion of a kallikrein polypeptide that is capable of modulating a function mediated by the kallikrein polypeptide.

An aptamer includes a DNA or RNA molecule that binds to polynucleotides and polypeptides. An aptamer that binds to a polypeptide (or binding domain) of a kallikrein polypeptide or a polynucleotide encoding a kallikrein polypeptide can be produced using conventional techniques, without undue experimentation. [For example, see the following publications describing *in vitro* selection of aptamers: Klug et al., Mol. Biol. Reports 20:97-107 (1994); Wallis et al., Chem. Biol. 2:543-552 (1995); Ellington, Curr. Biol. 4:427-429 (1994); Lato et al., Chem. Biol. 2:291-303 (1995); Conrad et al., Mol. Div. 1:69-78 (1995); and Uphoff et al., Curr. Opin. Struct. Biol. 6:281-287 (1996)].

In certain other preferred embodiments, the binding agent is an antibody.

In an aspect the present invention provides a diagnostic method for monitoring or diagnosing ovarian cancer in a subject by quantitating a plurality of kallikrein polypeptides, and optionally CA125, in a biological sample from the subject comprising reacting the sample with antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, which are directly or indirectly labelled with detectable substances, and detecting the detectable substances.

In an aspect of the invention, a method for detecting ovarian cancer is provided comprising:

- 15 (a) obtaining a sample suspected of containing a plurality of kallikrein polypeptides, and optionally CA125, wherein the kallikrein polypeptides comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11;
- (b) contacting the sample with antibodies that specifically bind to the plurality of kallikrein polypeptides, and optionally CA125, under conditions effective to bind the antibodies and form complexes;
- 20 (c) measuring the amount of kallikrein polypeptides, and optionally CA125, present in the sample by quantitating the amount of the complexes; and
- (d) comparing the amount of kallikrein polypeptides, and optionally CA125, present in the samples with the amount of polypeptides in a control, wherein a change or significant difference in the amount of polypeptides in the sample compared with the amount in the control is indicative of ovarian cancer.
- 25

In an embodiment, the invention contemplates a method for monitoring the progression of ovarian cancer in an individual, comprising:

- 30 (a) contacting antibodies which bind to each of a plurality of kallikrein polypeptides, and optionally CA125, with a sample from the individual so as to form binary complexes comprising each of the antibodies and polypeptides in the sample;
- (b) determining or detecting the presence or amount of complex formation in the sample;
- (c) repeating steps (a) and (b) at a point later in time; and
- 35 (d) comparing the result of step (b) with the result of step (c), wherein a difference in the amount of complex formation is indicative of the stage and/or progression of the ovarian cancer in said individual.

The amount of complexes may also be compared to a value representative of the amount of the complexes from an individual not at risk of, or afflicted with, ovarian cancer at different stages.

Thus, antibodies specifically reactive with each of a plurality of kallikrein polypeptides, and CA125, or derivatives, such as enzyme conjugates or labeled derivatives, may be used to detect a plurality of kallikrein polypeptides, and optionally CA125, in various samples (e.g. biological materials). They may be used as diagnostic or prognostic reagents and they may be used to detect abnormalities in the levels of expression of a plurality of kallikrein polypeptides, and optionally CA125, or abnormalities in the structure, and/or temporal, tissue, cellular, or subcellular location of a plurality of kallikrein polypeptides, and optionally CA125. Antibodies may also be used to screen potentially therapeutic compounds *in vitro* to determine their effects on ovarian cancer involving a plurality of kallikrein polypeptides, and optionally CA125, and other conditions. *In vitro* immunoassays may also be used to assess or monitor the efficacy of particular therapies.

Antibodies may be used in any known immunoassays that rely on the binding interaction between antigenic determinants of a plurality of kallikrein polypeptides, and optionally CA125, and the antibodies. Examples of such assays are radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence, immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. These terms are well understood by those skilled in the art. A person skilled in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

In particular, the antibodies may be used in immunohistochemical analyses, for example, at the cellular and sub-subcellular level, to detect a plurality of kallikrein polypeptides, and optionally CA125, to localize them to particular ovarian tumor cells and tissues, and to specific subcellular locations, and to quantitate the level of expression.

Antibodies for use in the present invention include monoclonal or polyclonal antibodies, immunologically active fragments (e.g. a Fab or (Fab)<sub>2</sub> fragments), antibody heavy chains, humanized antibodies, antibody light chains, genetically engineered single chain F<sub>v</sub> molecules (Ladner et al, U.S. Pat. No. 4,946,778), chimeric antibodies, for example, antibodies which contain the binding specificity of murine antibodies, but in which the remaining portions are of human origin, or derivatives, such as enzyme conjugates or labeled derivatives.

Antibodies including monoclonal and polyclonal antibodies, fragments and chimeras, may be prepared using methods known to those skilled in the art. Isolated native or recombinant kallikrein polypeptides or CA125 may be utilized to prepare antibodies. See, for example, Kohler et al. (1975) Nature 256:495-497; Kozbor et al. (1985) J. Immunol Methods 81:31-42; Cote et al. (1983) Proc Natl Acad Sci 80:2026-2030; and Cole et al. (1984) Mol Cell Biol 62:109-120 for the preparation of monoclonal antibodies; Huse et al. (1989) Science 246:1275-1281 for the preparation of monoclonal Fab fragments; and, Pound (1998) Immunochemical Protocols, Humana Press, Totowa, N.J for the preparation of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies. The antibodies specific for kallikrein polypeptides or CA125 used in the methods of the invention may also be obtained from scientific or commercial sources.

In an embodiment of the invention, antibodies are reactive against kallikrein polypeptides or CA125 if they bind with a K<sub>a</sub> of greater than or equal to 10<sup>-7</sup> M.

Antibodies that bind to kallikrein polypeptides or CA125 may be labelled with a detectable

substance and localised in biological samples based upon the presence of the detectable substance. Examples of detectable substances include, but are not limited to, the following: radioisotopes (e.g.,  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{35}\text{S}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ), fluorescent labels (e.g., FITC, rhodamine, lanthanide phosphors), luminescent labels such as luminol, enzymatic labels (e.g., horseradish peroxidase, beta-galactosidase, luciferase, alkaline phosphatase, acetylcholinesterase), biotinyl groups (which can be detected by marked avidin e.g., streptavidin containing a fluorescent marker or enzymatic activity that can be detected by optical or colorimetric methods), and predetermined polypeptide epitopes recognized by a secondary reporter (e.g., leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags). In some embodiments, labels are attached via spacer arms of various lengths to reduce potential steric hindrance. Antibodies may also be coupled to electron dense substances, such as ferritin or colloidal gold, which are readily visualised by electron microscopy.

Indirect methods may also be employed in which the primary antigen-antibody reaction is amplified by the introduction of a second antibody, having specificity for the antibody reactive against a kallikrein polypeptide or CA125. The second antibody may be labeled with a detectable substance to detect the primary antigen-antibody reaction. By way of example, if the antibody having specificity against a kallikrein polypeptide is a rabbit IgG antibody, the second antibody may be goat anti-rabbit gamma-globulin labelled with a detectable substance as described herein.

Methods for conjugating or labelling the antibodies discussed above may be readily accomplished by one of ordinary skill in the art. (See for example Inman, *Methods In Enzymology*, Vol. 34, *Affinity Techniques, Enzyme Purification: Part B*, Jakoby and Wichek (eds.), Academic Press, New York, p. 30, 1974; and Wilchek and Bayer, "The Avidin-Biotin Complex in Bioanalytical Applications," *Anal. Biochem.* 171:1-32, 1988 re methods for conjugating or labelling the antibodies with enzyme or ligand binding partner).

Cytochemical techniques known in the art for localizing antigens using light and electron microscopy may be used to detect a plurality of kallikrein polypeptides, and optionally CA125. Generally, antibodies may be labeled with detectable substances and kallikrein polypeptides, and optionally CA125, may be localised in tissues and cells based upon the presence of the detectable substance.

In the context of the methods of the invention, the sample, binding agents (e.g. antibodies) for a plurality of kallikrein polypeptides, and CA125 may be immobilized on a carrier or support. Examples of suitable carriers or supports are agarose, cellulose, nitrocellulose, dextran, Sephadex, Sepharose, liposomes, carboxymethyl cellulose, polyacrylamides, polystyrene, gabbros, filter paper, magnetite, ion-exchange resin, plastic film, plastic tube, glass, polyamine-methyl vinyl-ether-maleic acid copolymer, amino acid copolymer, ethylene-maleic acid copolymer, nylon, silk, etc. The support material may have any possible configuration including spherical (e.g. bead), cylindrical (e.g. inside surface of a test tube or well, or the external surface of a rod), or flat (e.g. sheet, test strip). Thus, the carrier may be in the shape of, for example, a tube, test plate, well, beads, disc, sphere, etc. The immobilized material may be prepared by reacting the material with a suitable insoluble carrier using known chemical or physical methods, for example, cyanogen bromide coupling. Binding agents (e.g. antibodies) may be indirectly immobilized using second binding agents specific for the first binding agent. For example, mouse antibodies specific for a kallikrein polypeptide may



be immobilized using sheep anti-mouse IgG Fc fragment specific antibody coated on the carrier or support.

Where radioactive labels are used as a detectable substance, a plurality of kallikrein polypeptides, and optionally CA125, may be localized by radioautography. The results of radioautography may be quantitated by determining the density of particles in the radioautographs by various optical methods, or by counting the grains.

Time-resolved fluorometry may be used to detect a signal. For example, the method described in Christopoulos TK and Diamandis EP Anal Chem 1992:64:342-346 may be used with a conventional time-resolved fluorometer.

Therefore, in accordance with an embodiment of the invention, a method is provided wherein antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125, are labelled with enzymes, substrates for the enzymes are added wherein the substrates are selected so that the substrates, or a reaction product of the enzymes and substrates, form fluorescent complexes with lanthanide metals. Lanthanide metals are added and the plurality of kallikrein polypeptides, and optionally CA125, are quantitated in the sample by measuring fluorescence of the fluorescent complexes. Antibodies specific for CA125 and each of a plurality of kallikrein polypeptides may be directly or indirectly labelled with enzymes. Enzymes are selected based on the ability of a substrate of the enzyme, or a reaction product of the enzyme and substrate, to complex with lanthanide metals such as europium and terbium. Examples of suitable enzymes include alkaline phosphatase and  $\beta$ -galactosidase.

Examples of enzymes and substrates for enzymes that provide such fluorescent complexes are described in U.S. Patent No. 5,312,922 to Diamandis. By way of example, when the antibody is directly or indirectly labelled with alkaline phosphatase the substrate employed in the method may be 4-methylumbelliferyl phosphate, 5-fluorosalicyl phosphate, or diflunisal phosphate. The fluorescence intensity of the complexes is typically measured using a time-resolved fluorometer e.g. a CyberFluor 615 Imunoanalyzer (Nordion International, Kanata, Ontario).

Antibodies specific for a plurality of kallikrein polypeptides and CA125 may also be indirectly labelled with enzymes. For example, an antibody may be conjugated to one partner of a ligand binding pair, and the enzyme may be coupled to the other partner of the ligand binding pair. Representative examples include avidin-biotin, and riboflavin-riboflavin binding protein. In another embodiment, antibodies specific for the anti-kallikrein antibodies or anti-CA125 antibodies are labeled with an enzyme.

In accordance with an embodiment, the present invention provides means for determining a plurality of kallikrein polypeptides, and optionally CA125, in a sample, in particular a serum sample, by measuring a plurality of kallikrein polypeptides, and optionally CA125, by immunoassay. It will be evident to a skilled artisan that a variety of immunoassay methods can be used to measure a plurality of kallikrein polypeptides and CA125 in serum. In general, an immunoassay method may be competitive or noncompetitive. Competitive methods typically employ immobilized or immobilizable antibodies to each of a plurality of kallikrein polypeptides, and optionally CA125, and a labeled form of each of a plurality of kallikrein polypeptides, and optionally CA125. Kallikrein polypeptides and CA125 and labeled kallikrein polypeptides and CA125 compete for binding to anti-kallikrein antibodies and anti-CA125 antibodies. After separation of the resulting labeled kallikrein polypeptides and CA125 that have become bound to anti-

kallikrein polypeptides and anti- CA125 (bound fraction) from that which has remained unbound (unbound fraction), the amount of the label in either bound or unbound fraction is measured and may be correlated with the amount of kallikrein polypeptides, and optionally CA125, in the test sample in any conventional manner, e.g., by comparison to a standard curve.

5 In an aspect, a non-competitive method is used for the determination of a plurality of kallikrein polypeptides, and optionally CA125, with the most common method being the "sandwich" method. In this assay, two types of antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125 are employed. One type of antibody is directly or indirectly labeled (sometimes referred to as the "detection antibody") and the other is immobilized or immobilizable (sometimes referred to as the "capture antibody"). The capture and detection antibodies can be contacted simultaneously or sequentially with a test sample. Sequential methods can be accomplished by incubating capture antibodies with the sample, and adding the detection antibodies at a predetermined time thereafter (sometimes referred to as the "forward" method); or the detection antibodies can be incubated with the sample first and then the capture antibodies added (sometimes referred to as the "reverse" method). After the necessary incubation(s) have occurred, to  
10 complete the assay, the capture antibodies are separated from the liquid test mixture, and labels are measured in at least a portion of the separated capture antibody phase or the remainder of the liquid test mixture. Generally the labels are measured in the capture antibody phase since it comprises kallikrein polypeptides, and optionally CA125, bound by ("sandwiched" between) the capture and detection antibodies. In an embodiment, the label may be measured without separating the capture antibodies and liquid test mixture.

20 In a typical two-site immunometric assay for a plurality of kallikrein polypeptides, and optionally CA125, one or both of the capture and detection antibodies are polyclonal antibodies or one or both of the capture and detection antibodies are monoclonal antibodies (i.e. polyclonal/polyclonal, monoclonal/monoclonal, or monoclonal/polyclonal). The labels used with the detection antibodies can be selected from any of those known conventionally in the art. The labels may be an enzyme or a  
25 chemiluminescent moiety, but it can also be a radioactive isotope, a fluorophor, a detectable ligand (e.g., detectable by a secondary binding by a labeled binding partner for the ligand), and the like. Preferably antibodies are labelled with enzymes which are detected by adding substrates that are selected so that a reaction product of the enzymes and substrates forms fluorescent complexes. Capture antibodies may be selected so that they provide a means for being separated from the remainder of the test mixture.  
30 Accordingly, the capture antibodies can be introduced to the assay in an already immobilized or insoluble form, or can be in an immobilizable form, that is, a form which enables immobilization to be accomplished subsequent to introduction of the capture antibodies to the assay. An immobilized capture antibody may comprise an antibody covalently or noncovalently attached to a solid phase such as a magnetic particle, a latex particle, a microtiter plate well, a bead, a cuvette, or other reaction vessel. An example of an  
35 immobilizable capture antibody is antibody which has been chemically modified with a ligand moiety, e.g., a hapten, biotin, or the like, and which can be subsequently immobilized by contact with an immobilized form of a binding partner for the ligand, e.g., an antibody, avidin, or the like. In an embodiment, a capture antibody may be immobilized using a species specific antibody for the capture antibody that is bound to the solid phase.

- 21 -

A particular sandwich immunoassay method of the invention employs two types of antibodies, first antibodies are reactive against each of a plurality of kallikrein polypeptides, and optionally CA125, and second antibodies having specificity against antibodies reactive against each of a plurality of kallikrein polypeptides, and optionally CA125, labelled with enzymatic labels, and fluorogenic substrates for the enzymes. An enzyme may be alkaline phosphatase (ALP) and the substrate is 5-fluorosalicyl phosphate. ALP cleaves phosphate out of the fluorogenic substrate, 5-fluorosalicyl phosphate, to produce 5-fluorosalicic acid (FSA). 5-Fluorosalicic acid can then form a highly fluorescent ternary complex of the form FSA-Tb(3+)-EDTA, which can be quantified by measuring the Tb3+ fluorescence in a time-resolved mode. Fluorescence intensity is measured using a time-resolved fluorometer as described herein.

The above-described immunoassay methods and formats are intended to be exemplary and are not limiting.

#### Computer Systems

Computer readable media comprising a plurality of kallikrein markers, and optionally CA125, is also provided. "Computer readable media" refers to any medium that can be read and accessed directly by a computer, including but not limited to magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. Thus, the invention contemplates computer readable medium having recorded thereon markers identified for patients and controls.

"Recorded" refers to a process for storing information on computer readable medium. The skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising information on a plurality of kallikrein markers, and optionally CA125.

A variety of data processor programs and formats can be used to store information on a plurality of kallikrein markers, and optionally CA125, on computer readable medium. For example, the information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. Any number of dataprocessor structuring formats (e.g., text file or database) may be adapted in order to obtain computer readable medium having recorded thereon the marker information.

By providing the marker information in computer readable form, one can routinely access the information for a variety of purposes. For example, one skilled in the art can use the information in computer readable form to compare marker information obtained during or following therapy with the information stored within the data storage means.

The invention provides a medium for holding instructions for performing a method for determining whether a patient has ovarian cancer or a pre-disposition to ovarian cancer, comprising determining the presence or absence of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and based on the presence or absence of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, determining whether the patient has ovarian cancer or a pre-disposition to

- 22 -

ovarian cancer, and optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention also provides in an electronic system and/or in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, comprising determining the presence or absence of a plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, and based on the presence or absence of the plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer, and optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention further provides in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, optionally CA125 and/or polynucleotides encoding same, comprising: (a) receiving phenotypic information on the subject and information on a plurality of kallikrein markers, optionally CA125 and/or polynucleotides encoding same associated with samples from the subject; (b) acquiring information from the network corresponding to the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same; and (c) based on the phenotypic information and information on the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer; and (d) optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention still further provides a system for identifying selected records that identify an ovarian cancer cell. A system of the invention generally comprises a digital computer; a database server coupled to the computer; a database coupled to the database server having data stored therein, the data comprising records of data comprising a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of records which match the desired selection criteria.

In an aspect of the invention a method is provided for detecting an ovarian cancer cell using a computer having a processor, memory, display, and input/output devices, the method comprising the steps of:

- (a) creating records of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, isolated from a sample suspected of containing an ovarian cancer cell;
- (b) providing a database comprising records of data comprising a plurality of kallikrein markers, optionally CA125, wherein the markers are kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and/or comprising polynucleotides encoding same; and
- (c) using a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of records of step (a) which provide a match of the desired selection criteria of the database of step (b) the presence of a match being a

- 23 -

positive indication that the markers of step (a) have been isolated from a cell that is an ovarian cancer cell.

The invention contemplates a business method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, comprising: (a) receiving phenotypic information on the subject and information on a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, associated with samples from the subject; (b) acquiring information from a network corresponding to the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same; and (c) based on the phenotypic information, information on a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and acquired information, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer; and (d) optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

#### Imaging Methods

Antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125, may also be used in imaging methodologies in the management of ovarian cancer. The invention provides a method for imaging tumors associated with a plurality of kallikrein polypeptides, and optionally CA125.

In an embodiment the method is an *in vivo* method and a subject or patient is administered imaging agents that carry imaging labels and are capable of targeting or binding to each of a plurality of kallikrein polypeptides, and optionally CA125. In the method each imaging agent is labeled so that it can be distinguished during the imaging. The imaging agents are allowed to incubate *in vivo* and bind to the plurality of kallikrein polypeptides, and optionally CA125, associated with ovarian tumors. The presence of label is localized to the ovarian cancer, and the localized label is detected using imaging devices known to those skilled in the art.

The imaging agents may be antibodies or chemical entities that recognize the plurality of kallikrein polypeptides, and optionally CA125. In an aspect of the invention an imaging agent is a polyclonal antibody or monoclonal antibody, or fragments thereof, or constructs thereof including but not limited to, single chain antibodies, bifunctional antibodies, molecular recognition units, and peptides or entities that mimic peptides. The antibodies specific for kallikrein polypeptides and CA125 used in the methods of the invention may be obtained from scientific or commercial sources, or isolated native or recombinant kallikrein and CA125 polypeptides may be utilized to prepare antibodies etc as described herein.

An imaging agent may be a peptide that mimics the epitope for an antibody specific for kallikrein polypeptide or CA125 and binds to kallikrein polypeptide or CA125. The peptide may be produced on a commercial synthesizer using conventional solid phase chemistry. By way of example, a peptide may be prepared that includes either tyrosine, lysine, or phenylalanine to which  $N_2S_2$  chelate is complexed (See U.S. Patent No. 4,897,255). The anti-kallikrein peptide conjugate is then combined with a radiolabel (e.g. sodium  $^{99m}Tc$  pertechnetate or sodium  $^{188}Re$  perrhenate) and it may be used to locate a tumor producing a plurality of kallikrein polypeptides, and optionally CA125.

Imaging agents carry labels to image the plurality of kallikrein polypeptides and CA125. Agents may be labelled for use in radionuclide imaging. In particular, agents may be directly or indirectly labelled

with a radioisotope. Examples of radioisotopes that may be used in the present invention are the following:  
<sup>277</sup>Ac, <sup>211</sup>At, <sup>128</sup>Ba, <sup>131</sup>Ba, <sup>7</sup>Be, <sup>204</sup>Bi, <sup>205</sup>Bi, <sup>206</sup>Bi, <sup>76</sup>Br, <sup>77</sup>Br, <sup>82</sup>Br, <sup>109</sup>Cd, <sup>47</sup>Ca, <sup>11</sup>C, <sup>14</sup>C, <sup>36</sup>Cl, <sup>48</sup>Cr, <sup>51</sup>Cr, <sup>62</sup>Cu,  
<sup>64</sup>Cu, <sup>67</sup>Cu, <sup>165</sup>Dy, <sup>155</sup>Eu, <sup>18</sup>F, <sup>153</sup>Gd, <sup>66</sup>Ga, <sup>67</sup>Ga, <sup>68</sup>Ga, <sup>72</sup>Ga, <sup>198</sup>Au, <sup>3</sup>H, <sup>166</sup>Ho, <sup>111</sup>In, <sup>113m</sup>In, <sup>115m</sup>In, <sup>123</sup>I, <sup>125</sup>I, <sup>131</sup>I,  
<sup>189</sup>Ir, <sup>191m</sup>Ir, <sup>192</sup>Ir, <sup>194</sup>Ir, <sup>52</sup>Fe, <sup>55</sup>Fe, <sup>59</sup>Fe, <sup>177</sup>Lu, <sup>15</sup>O, <sup>191m-191</sup>Os, <sup>109</sup>Pd, <sup>32</sup>P, <sup>33</sup>P, <sup>42</sup>K, <sup>226</sup>Ra, <sup>186</sup>Re, <sup>188</sup>Re, <sup>82m</sup>Rb,  
5 <sup>153</sup>Sm, <sup>46</sup>Sc, <sup>47</sup>Sc, <sup>72</sup>Se, <sup>75</sup>Se, <sup>105</sup>Ag, <sup>22</sup>Na, <sup>24</sup>Na, <sup>89</sup>Sr, <sup>35</sup>S, <sup>38</sup>S, <sup>177</sup>Ta, <sup>96</sup>Tc, <sup>99m</sup>Tc, <sup>201</sup>Tl, <sup>202</sup>Tl, <sup>113</sup>Sn, <sup>117m</sup>Sn,  
<sup>121</sup>Sn, <sup>166</sup>Yb, <sup>169</sup>Yb, <sup>175</sup>Yb, <sup>88</sup>Y, <sup>90</sup>Y, <sup>62</sup>Zn and <sup>65</sup>Zn. Preferably the radioisotope is <sup>131</sup>I, <sup>125</sup>I, <sup>123</sup>I, <sup>111</sup>I, <sup>99m</sup>Tc,  
<sup>90</sup>Y, <sup>186</sup>Re, <sup>188</sup>Re, <sup>32</sup>P, <sup>153</sup>Sm, <sup>67</sup>Ga, <sup>201</sup>Tl <sup>77</sup>Br, or <sup>18</sup>F, and it is imaged with a photoscanning device.

Procedures for labeling biological agents with the radioactive isotopes are generally known in the art. U.S. Pat. No. 4,302,438 describes tritium labeling procedures. Procedures for iodinating, tritium labeling,  
10 and <sup>35</sup>S labeling especially adapted for murine monoclonal antibodies are described by Goding, J. W. (supra, pp 124-126) and the references cited therein. Other procedures for iodinating biological agents, such as antibodies, binding portions thereof, probes, or ligands, are described in the scientific literature (see Hunter and Greenwood, Nature 144:945 (1962), David et al., Biochemistry 13:1014-1021 (1974), and U.S. Pat. Nos. 3,867,517 and 4,376,110). Iodinating procedures for agents are described by Greenwood, F. et al., Biochem.  
15 J. 89:114-123 (1963); Marchalonis, J., Biochem. J. 113:299-305 (1969); and Morrison, M. et al., Immunochemistry, 289-297 (1971). <sup>99m</sup>Tc-labeling procedures are described by Rhodes, B. et al. in Burchiel, S. et al. (eds.), Tumor Imaging: The Radioimmunochemical Detection of Cancer, New York: Masson 111-123 (1982) and the references cited therein. Labelling of antibodies or fragments with technetium-99m are also described for example in U.S. Pat. No. 5,317,091, U.S. Pat. No. 4,478,815, U.S.  
20 Pat. No. 4,478,818, U.S. Pat. No. 4,472,371, U.S. Pat. No. Re 32,417, and U.S. Pat. No. 4,311,688. Procedures suitable for <sup>111</sup>In-labeling biological agents are described by Hnatowich, D. J. et al., J. Immul. Methods, 65:147-157 (1983), Hnatowich, D. et al., J. Applied Radiation, 35:554-557 (1984), and Buckley, R. G. et al., F.E.B.S. 166:202-204 (1984).

An imaging agent may also be labeled with a paramagnetic isotope for purposes of an *in vivo*  
25 method of the invention. Examples of elements that are useful in magnetic resonance imaging include gadolinium, terbium, tin, iron, or isotopes thereof. (See, for example, Schaefer et al., (1989) JACC 14, 472-480; Shreve et al., (1986) Magn. Reson. Med. 3, 336-340; Wolf, G L., (1984) Physiol. Chem. Phys. Med. NMR 16, 93-95; Wesbey et al., (1984) Physiol. Chem. Phys. Med. NMR 16, 145-155; Runge et al., (1984) Invest. Radiol. 19, 408-415 for discussions on *in vivo* nuclear magnetic resonance imaging.)

30 In the case of radiolabeled agents, the agents may be administered to the patient, localized to the tumor having a plurality of kallikrein polypeptides, and optionally CA125, with which the agents bind, and detected or "imaged" *in vivo* using known techniques such as radionuclear scanning using, for example, a gamma camera or emission tomography. [See for example, A. R. Bradwell et al., "Developments in Antibody Imaging", Monoclonal Antibodies for Cancer Detection and Therapy, R. W. Baldwin et al., (eds.),  
35 pp. 65-85 (Academic Press 1985)]. A positron emission transaxial tomography scanner, such as designated Pet VI located at Brookhaven National Laboratory, can also be used where the radiolabel emits positrons (e.g., <sup>11</sup>C, <sup>18</sup>F, <sup>15</sup>O, and <sup>13</sup>N).

Whole body imaging techniques using radioisotope labeled agents can be used for locating both primary tumors and tumors which have metastasized. Antibodies specific for a plurality of kallikrein

polypeptides, and optionally CA125, or fragments thereof having the same epitope specificity, are bound to a suitable radioisotope, or a combination thereof, and administered parenterally. For ovarian cancer, administration preferably is intravenous. The bio-distribution of the labels can be monitored by scintigraphy, and accumulations of the labels can be related to the presence of ovarian cancer cells. Whole body imaging techniques are described in U.S. Pat. Nos. 4,036,945 and 4,311,688. Other examples of agents useful for diagnosis and therapeutic use that can be coupled to antibodies and antibody fragments include metallothionein and fragments (see, U.S. Pat. No. 4,732,864). These agents are useful in diagnosis, staging and visualization of cancer, in particular ovarian cancer, so that surgical and/or radiation treatment protocols can be used more efficiently.

#### 10 Screening Methods

The invention also contemplates methods for evaluating test agents or compounds for their ability to inhibit ovarian cancer or potentially contribute to ovarian cancer. Test agents and compounds include but are not limited to peptides such as soluble peptides including Ig-tailed fusion peptides, members of random peptide libraries and combinatorial chemistry-derived molecular libraries made of D- and/or L-configuration amino acids, phosphopeptides (including members of random or partially degenerate, directed phosphopeptide libraries), antibodies [e.g. polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, single chain antibodies, fragments, (e.g. Fab, F(ab)<sub>2</sub>, and Fab expression library fragments, and epitope-binding fragments thereof)], nucleic acids (e.g. antisense, interference RNA) and small organic or inorganic molecules. The agents or compounds may be endogenous physiological compounds or natural or synthetic compounds.

The invention also provides a method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, the method comprising comparing:

- (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from a patient and exposed to the test agent, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and
- (b) levels of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient, wherein the sample is not exposed to the test agent, wherein a significant difference in the levels of expression of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the first sample, relative to the second sample, is an indication that the test agent is potentially efficacious for inhibiting ovarian cancer in the patient.

The first and second samples may be portions of a single sample obtained from a patient or portions of pooled samples obtained from a patient.

In an aspect, the invention provides a method of selecting an agent for inhibiting ovarian cancer in a patient comprising:

- (a) obtaining a sample comprising cancer cells from the patient;
- (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents;
- (c) comparing a plurality of kallikrein markers, optionally CA125, and/or polynucleotides

- 26 -

encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and

- (d) selecting one of the test agents which alters the levels of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.

Still another aspect of the present invention provides a method of conducting a drug discovery business comprising:

- (a) providing one or more methods or assay systems for identifying agents that inhibit ovarian cancer in a patient;
- (b) conducting therapeutic profiling of agents identified in step (a), or further analogs thereof, for efficacy and toxicity in animals; and
- (c) formulating a pharmaceutical preparation including one or more agents identified in step (b) as having an acceptable therapeutic profile.

In certain embodiments, the subject method can also include a step of establishing a distribution system for distributing the pharmaceutical preparation for sale, and may optionally include establishing a sales group for marketing the pharmaceutical preparation.

The invention also contemplates a method of assessing the ovarian carcinogenic potential of a test compound comprising:

- (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and
- (b) comparing a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

A significant difference between the levels of the markers in the aliquot maintained in the presence of (or exposed to) the test compound relative to the aliquot maintained in the absence of the test compound, indicates that the test compound possesses ovarian carcinogenic potential.

#### **Kits**

The methods described herein may be performed by utilizing pre-packaged diagnostic kits comprising at least a plurality of kallikrein nucleic acids or binding agents (e.g. antibodies) or CA125 nucleic acids or binding agents described herein, which may be conveniently used, e.g., in clinical settings, to screen and diagnose patients, and to screen and identify those individuals afflicted with or exhibiting a predisposition to ovarian cancer.

Thus, the invention also contemplates kits for carrying out the methods of the invention. Such kits typically comprise two or more components required for performing a diagnostic assay. Components include but are not limited to compounds, reagents, containers, and/or equipment.

In an embodiment, a container with a kit comprises binding agents as described herein. By way of example, the kit may contain antibodies specific for a plurality of kallikrein polypeptides, and optionally



- 27 -

CA125, antibodies against the antibodies labelled with enzymes; and substrates for the enzymes. The kit may also contain microtiter plate wells, standards, assay diluent, wash buffer, adhesive plate covers, and/or instructions for carrying out a method of the invention using the kit.

5 In an aspect of the invention, the kit includes antibodies or antibody fragments which bind specifically to epitopes of each of a plurality of kallikrein polypeptides, and optionally CA125, and means for detecting binding of the antibodies to epitopes associated with tumor cells, either as concentrates (including lyophilized compositions), which may be further diluted prior to use or at the concentration of use, where the vials may include one or more dosages. Where the kits are intended for *in vivo* use, single dosages may be provided in sterilized containers, having the desired amount and concentration of agents.  
10 Containers that provide a formulation for direct use, usually do not require other reagents, as for example, where the kit contains radiolabelled antibody preparations for *in vivo* imaging.

A kit may be designed to detect the level of polynucleotides encoding kallikrein polypeptides, and optionally CA125 polynucleotides, in a sample. Such kits generally comprise oligonucleotide probes or primers, as described herein, that hybridize to a plurality of polynucleotides encoding kallikrein polypeptides  
15 and optionally CA125. Such oligonucleotides may be used, for example, within a PCR or hybridization procedure. Additional components that may be present within the kits include second oligonucleotides and/or diagnostic reagents to facilitate detection of a plurality polynucleotides encoding kallikrein polypeptides, and optionally CA125 polynucleotides.

The reagents suitable for applying the screening methods of the invention to evaluate compounds  
20 may be packaged into convenient kits described herein providing the necessary materials packaged into suitable containers.

#### Applications

Kallikrein polypeptides (in particular, kallikrein 5, 6, 10 and 11), optionally in combination with CA125, are targets for ovarian cancer immunotherapy. Such immunotherapeutic methods include the use of  
25 antibody therapy, *in vivo* vaccines, and *ex vivo* immunotherapy approaches.

In one aspect, the invention provides antibodies specific for a plurality of kallikrein polypeptides (for example, kallikreins 5, 6, 10 and 11) and optionally CA125, that may be used systemically to treat ovarian cancer. Preferably antibodies are used that target the tumor cells but not the surrounding non-tumor cells and tissue. Thus, the invention provides a method of treating a patient susceptible to, or having a cancer  
30 that expresses a plurality of kallikrein polypeptides, and optionally CA125, comprising administering to the patient an effective amount of antibodies that bind specifically to a plurality of kallikrein polypeptides, and optionally CA125. In another aspect, the invention provides a method of inhibiting the growth of tumor cells expressing a plurality of kallikrein polypeptides, and optionally CA125, comprising administering to a patient antibodies which bind specifically to the plurality of kallikrein polypeptides, and optionally CA125,  
35 in amounts effective to inhibit growth of the tumor cells. Antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may also be used in a method for selectively inhibiting the growth of, or killing a cell expressing a plurality of kallikrein polypeptides, and optionally CA125, comprising reacting antibody immunoconjugates or immunotoxins with the cell in an amount sufficient to inhibit the growth of, or kill the cell.

By way of example, unconjugated antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may be introduced into a patient such that the antibodies bind to cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125, and mediate growth inhibition of such cells (including the destruction thereof), and the tumor, by mechanisms which may include complement-mediated  
5 cytotoxicity, antibody-dependent cellular cytotoxicity, altering the physiologic function of a plurality of kallikrein polypeptides, and optionally CA125, and/or the inhibition of ligand binding or signal transduction pathways. In addition to unconjugated antibodies, antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, conjugated to therapeutic agents (e.g. immunoconjugates) may also be  
10 used therapeutically to deliver the agents directly to tumor cells expressing a plurality of kallikrein polypeptides, and optionally CA125, and thereby destroy the tumor. Examples of such agents include abrin, ricin A, *Pseudomonas* exotoxin, or diphtheria toxin, proteins such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, and biological response modifiers such as lymphokines, interleukin-1, interleukin-2, interleukin-6, granulocyte  
15 macrophage colony stimulating factor, granulocyte colony stimulating factor, or other growth factors.

Cancer immunotherapy using antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may utilize the various approaches that have been successfully employed for cancers, including but not limited to colon cancer (Arlen et al., 1998, Crit Rev Immunol 18: 133-138), multiple myeloma (Ozaki et al., 1997, Blood 90: 3179-3186; Tsunenati et al., 1997, Blood 90: 2437-2444), gastric cancer (Kasprzyk et al., 1992, Cancer Res 52: 2771-2776), B-cell lymphoma (Funakoshi et al., 1996, J  
20 Immunother Emphasis Tumor Immunol 19: 93-101), leukemia (Zhong et al., 1996, Leuk Res 20: 581-589), colorectal cancer (Moun et al., 1994, Cancer Res 54: 6160-6166); Velders et al., 1995, Cancer Res 55: 4398-4403), and breast cancer (Shepard et al., 1991, J Clin Immunol 11: 117-127).

In the practice of a method of the invention, antibodies specific for a plurality of kallikrein polypeptides, optionally in combination with antibodies specific for CA125, capable of inhibiting the growth  
25 of cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125, are administered in a therapeutically effective amount to cancer patients whose tumors express or overexpress a plurality of kallikrein polypeptides, and optionally CA125. The invention may provide a specific, effective and long-needed treatment for ovarian cancer. The antibody therapy methods of the invention may be combined with other therapies including chemotherapy and radiation.

30 Patients may be evaluated for the presence and levels of a plurality of kallikrein polypeptides, and optionally CA125, expression and overexpression in tumors, preferably using immunohistochemical assessments of tumor tissue, quantitative imaging as described herein, or other techniques capable of reliably indicating the presence and degree of expression of a plurality of kallikrein polypeptides, and optionally CA125. Immunohistochemical analysis of tumor biopsies or surgical specimens may be employed for this  
35 purpose.

Antibodies specific for a plurality of kallikrein polypeptides and CA125 useful in treating cancer include those that are capable of initiating a potent immune response against the tumor and those that are capable of direct cytotoxicity. In this regard, the antibodies may elicit tumor cell lysis by either complement-mediated or antibody-dependent cell cytotoxicity (ADCC) mechanisms, both of which require an intact Fc

- 29 -

portion of the immunoglobulin molecule for interaction with effector cell Fc receptor sites or complement proteins. In addition, antibodies specific for a plurality of kallikrein polypeptides and CA125 that exert a direct biological effect on tumor growth are useful in the practice of the invention. Such antibodies may not require the complete immunoglobulin to exert the effect. Potential mechanisms by which such directly cytotoxic antibodies may act include inhibition of cell growth, modulation of cellular differentiation, modulation of tumor angiogenesis factor profiles, and the induction of apoptosis. The mechanism by which a particular antibody exerts an anti-tumor effect may be evaluated using any number of *in vitro* assays designed to determine ADCC, antibody-dependent macrophage-mediated cytotoxicity (ADMMC), complement-mediated cell lysis, and others known in the art.

10 The anti-tumor activity of a combination of antibodies specific for a plurality of kallikrein polypeptides and optionally CA125, may be evaluated *in vivo* using a suitable animal model. Xenogenic cancer models, wherein human cancer explants or passaged xenograft tissues are introduced into immune compromised animals, such as nude or SCID mice, may be employed.

15 The methods of the invention contemplate the administration of combinations, or "cocktails" of different individual antibodies recognizing epitopes of a plurality of kallikrein polypeptides, and optionally CA125. Such cocktails may have certain advantages inasmuch as they contain antibodies that bind to different epitopes and/or exploit different effector mechanisms or combine directly cytotoxic antibodies with antibodies that rely on immune effector functionality. Such antibodies in combination may exhibit synergistic therapeutic effects. In addition, the administration of the antibodies may be combined with other therapeutic agents, including but not limited to chemotherapeutic agents, androgen-blockers, and immune modulators (e.g., IL2, GM-CSF). The antibodies may be administered in their "naked" or unconjugated form, or may have therapeutic agents conjugated to them.

20 The antibodies specific for a plurality of kallikrein polypeptides and optionally CA125, used in the practice of the method of the invention may be formulated into pharmaceutical compositions comprising a carrier suitable for the desired delivery method. Suitable carriers include any material which when combined with the antibodies retains the anti-tumor function of the antibodies and is non-reactive with the subject's immune systems. Examples include any of a number of standard pharmaceutical carriers such as sterile phosphate buffered saline solutions, bacteriostatic water, and the like (see, generally, Remington's Pharmaceutical Sciences 16.sup.th Edition, A. Osal., Ed., 1980).

30 Antibody formulations may be administered via any route capable of delivering the antibodies to the tumor site. Routes of administration include, but are not limited to, intravenous, intraperitoneal, intramuscular, intratumor, intradermal, and the like. Preferably, the route of administration is by intravenous injection. Antibody preparations may be lyophilized and stored as a sterile powder, preferably under vacuum, and then reconstituted in bacteriostatic water containing, for example, benzyl alcohol preservative, or in sterile water prior to injection.

35 Treatment will generally involve the repeated administration of the antibody preparation via an acceptable route of administration such as intravenous injection (IV), at an effective dose. Dosages will depend upon various factors generally appreciated by those of skill in the art, including the type of cancer and the severity, grade, or stage of the cancer, the binding affinity and half life of the antibodies used, the

degree of expression of a plurality of kallikrein polypeptides, and optionally CA125, in the patient, the extent of circulating kallikrein polypeptide antigens, and optionally CA125 antigens, the desired steady-state antibody concentration level, frequency of treatment, and the influence of any chemotherapeutic agents used in combination with a treatment method of the invention.

5       Daily doses may range from about 0.1 to 100 mg/kg. Doses in the range of 10-500 mg antibodies per week may be effective and well tolerated, although even higher weekly doses may be appropriate and/or well tolerated. A determining factor in defining the appropriate dose is the amount of antibodies necessary to be therapeutically effective in a particular context. Repeated administrations may be required to achieve tumor inhibition or regression. Direct administration of antibodies specific for a plurality of kallikrein polypeptides and optionally CA125 is also possible and may have advantages in certain situations.

10       Patients may be evaluated for a plurality of kallikrein polypeptides and optionally CA125, preferably in serum, in order to assist in the determination of the most effective dosing regimen and related factors. The assay methods described herein, or similar assays, may be used for quantitating circulating kallikrein polypeptide and optionally CA125 levels in patients prior to treatment. Such assays may also be used for monitoring throughout therapy, and may be useful to gauge therapeutic success in combination with evaluating other parameters, such as serum kallikrein polypeptides, and optionally CA125, levels.

15       The invention further provides vaccines formulated to contain a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof. The use in anti-cancer therapy of tumor antigens in a vaccine for generating humoral and cell-mediated immunity is well known and, for example, has been employed in prostate cancer using human PSMA and rodent PAP immunogens (Hodge et al., 1995, Int. J. Cancer 63: 231-237; Fong et al., 1997, J. Immunol. 159: 3113-3117). These methods can be practiced by employing a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof, or nucleic acids and recombinant vectors capable of expressing and appropriately presenting the kallikrein and optionally CA125, immunogens.

20       By way of example, viral gene delivery systems may be used to deliver nucleic acids encoding a plurality of kallikrein polypeptides, and optionally CA125. Various viral gene delivery systems which can be used in the practice of this aspect of the invention include, but are not limited to, vaccinia, fowlpox, canarypox, adenovirus, influenza, poliovirus, adeno-associated virus, lentivirus, and sindbus virus (Restifo, 1996, Curr. Opin. Immunol. 8: 658-663). Non-viral delivery systems may also be employed by using naked DNA encoding a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof introduced into the patient (e.g., intramuscularly) to induce an anti-tumor response.

25       Various *ex vivo* strategies may also be employed. One approach involves the use of cells to present kallikrein and optionally CA125 antigens to a patient's immune system. For example, autologous dendritic cells which express MHC class I and II, may be pulsed with a plurality of kallikrein polypeptides, and optionally CA125, or peptides thereof that are capable of binding to MHC molecules, to thereby stimulate ovarian cancer patients' immune systems (See, for example, Tjoa et al., 1996, Prostate 28: 65-69; Murphy et al., 1996, Prostate 29: 371-380).

30       Anti-idiotypic antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, can also be used in anti-cancer therapy as a vaccine for inducing an immune response to cells expressing the

polypeptides. The generation of anti-idiotypic antibodies is well known in the art and can readily be adapted to generate anti-idiotypic antibodies that mimic an epitope on a kallikrein polypeptide or CA125 (see, for example, Wagner et al., 1997, Hybridoma 16: 33-40; Foon et al., 1995, J Clin Invest 96: 334-342; Herlyn et al., 1996, Cancer Immunol Immunother 43: 65-76). Such antibodies can be used in anti-idiotypic therapy as presently practiced with other anti-idiotypic antibodies directed against tumor antigens.

Genetic immunization methods may be utilized to generate prophylactic or therapeutic humoral and cellular immune responses directed against cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125. Constructs comprising DNA encoding kallikrein and optionally CA125 polypeptides/immunogens and appropriate regulatory sequences may be injected directly into muscle or skin of an individual, such that the cells of the muscle or skin take-up the construct and express the encoded kallikrein or CA125 polypeptides/immunogens. The polypeptides/immunogens may be expressed as cell surface proteins or be secreted. Expression of the polypeptides/immunogens results in the generation of prophylactic or therapeutic humoral and cellular immunity against the cancer. Various prophylactic and therapeutic genetic immunization techniques known in the art may be used.

The invention further provides methods for inhibiting cellular activity (e.g., cell proliferation, activation, or propagation) of a cell expressing a plurality of kallikrein polypeptides, and optionally CA125. This method comprises reacting immunoconjugates of the invention (e.g., a heterogeneous or homogenous mixture) with the cell so that the kallikrein polypeptides, and optionally CA125, form complexes with the immunoconjugates. A subject with a neoplastic or preneoplastic condition can be treated when the inhibition of cellular activity results in cell death.

In another aspect, the invention provides methods for selectively inhibiting a cell expressing a plurality of kallikrein polypeptides, and optionally CA125, by reacting a combination of immunoconjugates of the invention with the cell in an amount sufficient to inhibit the cell. Amounts include those that are sufficient to kill the cell or sufficient to inhibit cell growth or proliferation.

Vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids, may be used to deliver nucleic acids encoding a plurality of kallikrein polypeptides, and optionally CA125, to a targeted organ, tissue, or cell population. Methods well known to those skilled in the art may be used to construct recombinant vectors that will express antisense nucleic acid molecules for kallikrein polypeptides and CA125. (See, for example, the techniques described in Sambrook et al (supra) and Ausubel et al (supra)).

Genes encoding a plurality of kallikrein polypeptides, and optionally CA125, can be turned off by transfecting a cell or tissue with vectors that express high levels of a desired kallikrein or CA125 polypeptide-encoding fragments. Such constructs can inundate cells with untranslatable sense or antisense sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until all copies are disabled by endogenous nucleases.

Modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA or PNA, to the regulatory regions of genes encoding kallikrein polypeptides, and optionally CA125, i.e., the promoters, enhancers, and introns. Preferably, oligonucleotides are derived from the transcription initiation site, eg, between -10 and +10 regions of the leader sequence. The antisense molecules may also be designed

so that they block translation of mRNA by preventing the transcript from binding to ribosomes. Inhibition may also be achieved using "triple helix" base-pairing methodology. Triple helix pairing compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Therapeutic advances using triplex DNA were reviewed by Gee J E et al (In: Huber B E and B I Carr (1994) Molecular and Immunologic Approaches, Futura Publishing Co, Mt Kisco N.Y.).

Ribozymes are enzymatic RNA molecules that catalyze the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. The invention therefore contemplates engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding a plurality of kallikrein polypeptides, and optionally CA125.

Specific ribozyme cleavage sites within any potential RNA target may initially be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once the sites are identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be determined by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Methods for introducing vectors into cells or tissues include those methods discussed herein and which are suitable for *in vivo*, *in vitro* and *ex vivo* therapy. For *ex vivo* therapy, vectors may be introduced into stem cells obtained from a patient and clonally propagated for autologous transplant into the same patient (See U.S. Pat. Nos. 5,399,493 and 5,437,994). Delivery by transfection and by liposome are well known in the art.

Kallikrein polypeptides, optionally CA125 polypeptide, and/or polynucleotides encoding the polypeptides, and fragments thereof, antibodies and/or agents identified using a method of the invention, or combinations thereof, may be used in the treatment of ovarian cancer or diseases, conditions or syndromes associated with ovarian cancer, in a subject. A combination of kallikrein polypeptides and/or polynucleotides encoding the kallikreins (e.g. kallikreins 7 and 8) and inhibitors (antisense, antibodies, or agents) of other kallikreins (e.g. kallikreins 5, 6, 10 and 11) and/or CA125 may be used in a prognostic or therapeutic method of the invention. The polypeptides, polynucleotides, and agents may be formulated into compositions for administration to subjects suffering from ovarian cancer. Therefore, the present invention also relates to a composition comprising a plurality of kallikrein polypeptides and optionally CA125, or nucleic acids encoding the polypeptides, or a fragment thereof, or an agent identified using a method of the invention, and a pharmaceutically acceptable carrier, excipient or diluent. A method for treating or preventing ovarian cancer in a subject is also provided comprising administering to a patient in need thereof, a plurality of kallikrein polypeptides and optionally CA125, or nucleic acids encoding the polypeptides, an agent identified in accordance with a method of the invention, and/or a composition of the invention.

The active substance may be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active substance may be coated in a material to

protect the substance from the action of enzymes, acids and other natural conditions that may inactivate the substance.

The compositions described herein can be prepared by per se known methods for the preparation of pharmaceutically acceptable compositions which can be administered to subjects, such that an effective quantity of the active substance is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example, in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). On this basis, the compositions include, albeit not exclusively, solutions of the active substances in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable pH and iso-osmotic with the physiological fluids.

The compositions are indicated as therapeutic agents either alone or in conjunction with other therapeutic agents or other forms of treatment (e.g. chemotherapy or radiotherapy). The compositions of the invention may be administered concurrently, separately, or sequentially with other therapeutic agents or therapies.

The following non-limiting examples are illustrative of the present invention:

#### Example 1

To investigate the additional discriminatory value of the kallikreins to CA125 a logistic regression model was developed. Included in the study were serum samples from 39 ovarian cancer patients and 194 non-cancer females. The age of the patients was as follows: median = 59, range 32-82. The age of the controls was as follows: median = 46; range = 22-77. The model was adjusted for the following variables:  $f(x) = -0.29 \text{ hK5} + 0.12 \text{ hK6} - 0.65 \text{ hK7} - 0.6 \text{ hK8} + 1.09 \text{ hK10} + 0.98 \text{ hK11} + 0.057 \text{ CA125} - 0.62$ . For these data, the crude odds ratio and the 95% confidence interval were found to be 2.71 and 1.91-3.84 ( $p < 0.001$ ). The log likelihood scores for this multivariate logistic regression model, which incorporated the combined variables for each patient was calculated. From these data, by picking different thresholds for the regression function values, a ROC curve was devised which shows the added value of using kallikreins and CA125 together in a multivariate function. (AUC, 0.99; 95%CI, 0.96-1.00). (See Figure 8.) Statistically significant correlations between age and other studied variables were not observed.

#### Example 2

Statistically significant differences in serum kallikrein concentration was found between patient and control subjects for kallikreins hK5 ( $p < 0.0001$ ), hK7 ( $p = 0.007$ ), hK8 ( $p = 0.005$ ), hK10 ( $p = 0.0003$ ) and CA125 ( $p < 0.0001$ ) by the Mann-Whitney test. The diagnostic sensitivity (SENS) and specificity (SPEC) for each one of these markers were as follows (SENS/SPEC; both as %): 31/95 (hK5); 62/71 (hK7); 62/70 (hK10); 54/54 (hK11); 89/94 (CA125). When these data were combined in a logistic regression model, kallikreins 5 and 10 did not contribute to a great extent to the sensitivity and specificity of CA125. The area under the curve of CA125 alone (93%) improved by a further 1% when adding hK6, by 2% when adding hK11, 3% when adding hK7 and 5% when adding hK8. The combination of CA125 and hK8 resulted in an AUC of 98%.

Below is a summary of each marker and its ability to separate the cases and controls.

- 34 -

- hK5: high values associated with cancer  
 test+ is  $hK5 > 0.10$ , test- is  $hK5 \leq 0.10$   
 sensitivity=31%, specificity=95%,  
 AUC=.62,  $p(AUC)=.02$   
 5 Wilcoxon rank sum test has  $p < .0001$ .  
 Of the 233 persons analyzed, 207 have value zero for hK5 (27 cases, 180 controls).  
 Possible good marker
- hK6: high values associated with cancer  
 10 test+ is  $hK6 > 6.3$ , test- is  $hK6 \leq 6.3$   
 sensitivity=69%, specificity=40%,  
 AUC=.50,  $p(AUC)=1.00$   
 Wilcoxon rank sum test has  $p=.91$ .  
 Not a good marker
- hK7: low values associated with cancer  
 15 test+ is  $hK7 < 2.05$ , test- is  $hK7 \geq 2.05$   
 sensitivity=62%, specificity=71%,  
 AUC=.64,  $p(AUC)=.006$   
 20 Wilcoxon rank sum test has  $p=.007$ .  
 Possible good marker
- hK8: low values associated with cancer  
 25 test+ is  $hK8 < 13.0$ , test- is  $hK8 \geq 13.0$   
 sensitivity=72%, specificity=42%,  
 AUC=.64,  $p(AUC)=.006$   
 Wilcoxon rank sum test has  $p=.005$   
 Possible good marker
- hK10: high values associated with cancer  
 30 test+ is  $hK10 > 1.42$ , test- is  $hK10 \leq 1.42$   
 sensitivity=62%, specificity=70%,  
 AUC=.68,  $p(AUC)=.0004$   
 Wilcoxon rank sum test has  $p=.0003$ .  
 35 Best single kallikrein marker
- hK11: high values associated with cancer  
 40 test+ is  $hK11 > 0.14$ , test- is  $hK11 \leq 0.14$   
 sensitivity=54%, specificity=54%,  
 AUC=.58,  $p(AUC)=.12$   
 Wilcoxon rank sum test has  $p=.11$ .  
 Not a good marker
- CA125: high values associated with cancer  
 45 test+ is  $Ca125 > 34$ , test- is  $Ca125 \leq 34$   
 sensitivity=89%, specificity=94%,  
 AUC=.933,  $p(AUC) < .0001$   
 Wilcoxon rank sum test has  $p < .0001$ .  
 Good marker
- 50 After some further multivariate analysis of only the kallikrein markers, the combination of hK 7, 8, 10 and 11 was a preferred set. This combination was arrived at by looking at the incremental AUC as markers were combined. Below is a summary of all the models tried:
- 55 hK10 alone, AUC=.68  
 hK10+hK7: AUC=.88



- 35 -

hK10+hK7+hK8: AUC=.90  
 hK10+hK7+hK8+hK11: AUC=.925  
 Multivariate model of hK7, hK8, hK10, hK11, call it hK7\_8\_10\_11

- 5 hK7\_8\_10\_11:  
 Calculate  $SA = 2.00 - 1.49(hK7) - .34(hK8) + 1.16(hK10) + 3.50(hK11)$   
 high values associated with cancer  
 test+ is  $SA > 1.15$ , test- is  $SA \leq 1.15$   
 sensitivity=87%, specificity=89%,  
 10 AUC=.93,  $p(AUC) < .0001$   
 Wilcoxon rank sum test has  $p < .0001$ .  
 Good marker

The hK marker that added the most to CA125 was also investigated.

- 15 CA125 alone, AUC=.933  
 CA125+hK8: AUC=.978  
 Multivariate model of Ca125, hK8, call it Ca125\_hK8

- 20 Ca125\_hK8:  
 $SC = 1.71 + .086(Ca125) - .47(hK8)$   
 high values associated with cancer  
 test+ is  $SC > 2.52$ , test- is  $SC \leq 2.52$   
 sensitivity=97%, specificity=90%,  
 AUC=.978,  $p(AUC) < .0001$   
 25 Wilcoxon rank sum test has  $p < .0001$ .  
 Good marker

Below is a summary of the above analyses:

- a) The preferred kallikrein marker alone is hK10, AUC=.68  
 30 b) CA125 has an AUC of .933  
 c) The preferred combination of kallikrein markers increases the AUC up to .925, which is close to the CA125 AUC of .933  
 d) Adding a kallikrein marker to CA125 can increase the AUC up to .978
- 35 How does CA125 alone compare with the multivariate kallikrein model hK7\_8\_10\_11?  
 (based on 39 cases and 186 controls evaluated with CA125)

	Sensitivity	Specificity	misclassification
40 CA125	90%	94%	12FP, 4FN, total 16 pts misclassified
hK7_8_10_11	85%	89%	31FP, 4FN, total 35 pts misclassified
both positive	77%	100%	0FP, 9FN, total 9 pts misclassified
either positive	97%	82%	33FP, 1FN, total 34 pts misclassified

How does CA125 alone compare with the multivariate model of CA125 plus hK8?

- 45 (based on 39 cases and 186 controls evaluated with CA125)
- |           | Sensitivity | Specificity | misclassification                     |
|-----------|-------------|-------------|---------------------------------------|
| CA125     | 90%         | 94%         | 12FP, 4FN, total 16 pts misclassified |
| CA125_hK8 | 95%         | 91%         | 17FP, 2FN, total 19 pts misclassified |

Kallikrein markers approach CA125 in terms of AUC and sensitivity, but their specificity is not as high. Adding hK8 to CA125 improves sensitivity but specificity is lower than CA125 alone.

Summary

- 5           a)       The best kallikrein marker alone is hK10, area under the ROC curve (AUC) = .68.
- b)       CA125 has an AUC of .933. Adding a single kallikrein marker to CA125 can get the AUC up to .978. Adding hK8 to CA125 improves sensitivity but specificity is lower compared with CA125 alone.
- 10           c)       The best combination of kallikrein markers gets the AUC up to .925, which is close to the CA125 AUC of .933. Kallikrein markers approach CA125 in terms of AUC and sensitivity, but their specificity is lower.

15           The present invention is not to be limited in scope by the specific embodiments described herein, since such embodiments are intended as but single illustrations of one aspect of the invention and any functionally equivalent embodiments are within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

20           All publications, patents and patent applications referred to herein are incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety. All publications, patents and patent applications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the domains, cell lines, vectors, methodologies etc. which are reported therein

25           which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

            It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more

30           antibodies and equivalents thereof known to those skilled in the art, and so forth.

            Below full citations are set out for the references referred to in the specification.

Table 1

Kallikrein Polypeptide	Kallikrein Nucleic Acid Designation	GenBank Accession No.
Kallikrein 5	KLK5	AAD26429, AF135028, AF168768
Kallikrein 6	KLK6	AAB66483, AF013988 (CDS 174..881), AF149289 (CDS join 3567..3606, 4346..4502, 8122..8369, 9791..9927, 11805..11957) U62801 (CDS 246..980)
Kallikrein 7	KLK7	AAC37551, L33404 (CDS 16..777), AF166330 (CDS join 3237..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7517) (mRNA join(1756..1785, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..8265) /product="stratum corneum chymotryptic enzyme" /note="alternatively spliced" ; mRNA join (1756..1785, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7991) /note="alternatively spliced"; mRNA join (1821..1864, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..8265) /product="stratum corneum chymotryptic enzyme" /note="alternatively spliced"; mRNA join (1821..1864, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7991) /note="alternatively spliced"
Kallikrein 8	KLK8	BAA28673, AB009849 (CDS 35..817), AF095743 (CDS join 1035..1104, 1619..1778, 1944..2206, 4304..4437, 5974..6129, mRNA 500..670, 1027..1104, 1619..1778, 1944..2206, 4304..4437, 5974..6174), AB010780 (CDS join 1..39, 418..712, 878..>946), AF055982
Kallikrein 10	KLK10	AAC14266, AF055481 (CDS join 614..701, 2455..2635, 3589..3863, 4195..4328, 4793..4945, mRNA join 48..120, 605..701, 2455..2635, 3589..3863, 4195..4328, 4793..5474), NM_002776 (CDS 220..1050)
Kallikrein 11	KLK11	BAA33404, AAD47815, AB012917 (CDS 26..874), AF164623 (CDS 4224..4263, 5061..5217, 5545..5810, 6627..6763, 7158..7310) (mRNA join (2313..2398, 4189..4263, 5061..5217, 5545..5810, 6627..6763, 7158..7622)

- 38 -

Table 2

Descriptive statistics for hk5, hk6, hk7, hk8, hk10 and hk11 serum protein levels in controls and patients with ovarian cancer

	Mean	Standard Error	Median	Range	p value*
<u>hk5 (ng/ml)</u>					
Non cancer (N=194)	0.063	0.029	0.00	0.00-4.50	
Cancer (N=39)	0.48	0.18	0.00	0.00-5.70	
% Increase**	661%	--	--	--	<0.001
<u>hk6(ng/ml)</u>					
Non cancer (N=194)	6.96	0.18	6.60	1.60-15.30	
Cancer (N=39)	9.88	2.20	6.60	1.50-70.80	
% Increase**	42%	--	--	--	0.91
<u>hk7(ng/ml)</u>					
Non cancer (N=194)	2.60	0.071	2.67	0.30-6.00	
Cancer (N=39)	2.49	0.41	1.80	0.00-10.80	
% Decrease**	4%	--	33%	--	0.007
<u>hk8(ng/ml)</u>					
Non cancer (N=194)	11.74	0.27	11.70	2.40-22.20	
Cancer (N=39)	11.91	1.88	6.90	0.00-46.20	
% Decrease**	--	--	41%	--	0.005
<u>hk10(ng/ml)</u>					
Non cancer (N=194)	1.16	0.051	1.08	0.00-4.20	
Cancer (N=39)	6.51	2.46	1.59	0.27-90.0	
% Increase**	461%	--	40%	--	<0.001
<u>hk11(ng/ml)</u>					
Non cancer (N=194)	0.21	0.018	0.12	00-1.30	
Cancer (N=39)	0.79	0.21	0.18	0.00-5.52	
% Increase**	276%	--	50%		0.011

\* Calculated by the Mann Whitney test

\*\* Calculated by assuming that value in non-cancerous tissue is 100%

**Table 3**  
**Correlations between the studied variables in 194 non-cancer cases**

variable		hK5	hK6	hK7	hK8	hK10	hK11	CA125
<b>hK5</b>	$r_s$	1.000	0.034	-0.053	0.066	0.134	0.150	0.101
	$p$	.	0.642	0.462	0.359	0.062	0.037	0.172
<b>hK6</b>	$r_s$	0.034	1.000	0.114	0.298	0.191	0.120	-0.160
	$p$	0.642	.	0.115	0.000	0.008	0.097	0.029
<b>hK7</b>	$r_s$	-0.053	0.114	1.000	0.497	0.321	0.399	0.135
	$p$	0.462	0.115	.	0.000	0.000	0.000	0.066
<b>hK8</b>	$r_s$	0.066	0.298	0.497	1.000	0.263	0.396	0.048
	$p$	0.359	0.000	0.000	.	0.000	0.000	0.519
<b>hK10</b>	$r_s$	0.134	0.191	0.321	0.263	1.000	0.176	0.035
	$p$	0.062	0.008	0.000	0.000	.	0.014	0.638
<b>hK11</b>	$r_s$	0.150	0.120	0.399	0.396	0.176	1.000	0.125
	$p$	0.037	0.097	0.000	0.000	0.014	.	0.090
<b>CA125</b>	$r_s$	0.101	-0.160	0.135	0.048	0.035	0.125	1.000
	$p$	0.172	0.029	0.066	0.519	0.638	0.090	.

**Table 4**  
**Correlations between the studied variables in 39 ovarian cancer cases**

variable		hK5	hK6	hK7	hK8	hK10	hK11	CA125
<b>hK5</b>	$r_s$	1.000	0.475	0.553	0.554	0.618	0.584	0.507
	$p$		0.002	0.000	0.000	0.000	0.000	0.001
<b>hK6</b>	$r_s$	0.475	1.000	0.327	0.513	0.470	0.661	0.530
	$p$	0.002		0.042	0.001	0.003	0.000	0.001
<b>hK7</b>	$r_s$	0.553	0.327	1.000	0.695	0.690	0.748	0.262
	$p$	0.000	0.042		0.000	0.000	0.000	0.107
<b>hK8</b>	$r_s$	0.554	0.513	0.695	1.000	0.602	0.783	0.443
	$p$	0.000	0.001	0.000		0.000	0.000	0.005
<b>hK10</b>	$r_s$	0.618	0.470	0.690	0.602	1.000	0.706	0.548
	$p$	0.000	0.003	0.000	0.000		0.000	0.000
<b>hK11</b>	$r_s$	0.584	0.661	0.748	0.783	0.706	1.000	0.556
	$p$	0.000	0.000	0.000	0.000	0.000		0.000
<b>CA125</b>	$r_s$	0.507	0.530	0.262	0.443	0.548	0.556	1.000
	$p$	0.001	0.001	0.107	0.005	0.000	0.000	

Table 5

	CANCER				CONTROL				Wilcoxon		p(logistic)		AUC (se)		p(AUC)		CUTPOINT		SENSITIVITY		SPECIFICITY	
	n	mean	median	sd	min	max	n	mean	median	sd	min	max	p(mean)	p(median)								
HK5	39	0.48	0.00	0.00	1.14	0.00	194	0.06	0.00	0.00	0.40	0.00	0.03	<0.001	0.005	0.62 (.05)	0.0200	>0.10		31%		95% HK5
HK6	39	9.28	0.60	0.60	13.77	1.60	194	5.07	0.97	6.60	2.60	16.30	0.20	0.0001	0.0570	0.50 (.05)	1.0000	>6.30		69%		40% HK6
HK7	39	2.49	1.80	2.49	0.00	10.90	194	2.60	2.67	2.67	0.98	0.30	0.76	0.0070	0.6500	0.64 (.05)	0.0060	<2.05		62%		71% HK7
HK8	39	11.91	0.90	0.90	11.80	0.00	194	11.74	11.70	3.80	2.40	22.20	0.83	0.0050	0.8700	0.64 (.05)	0.0060	<13.0		72%		42% HK8
HK10	39	6.52	1.59	1.59	15.40	0.27	194	1.17	1.17	1.08	0.72	0.00	0.94	0.0003	0.0002	0.68 (.05)	0.0004	>1.42		62%		70% HK10
HK11	39	0.60	0.18	0.18	1.36	0.00	194	0.21	0.21	0.12	0.26	0.00	0.01	0.1100	0.0008	0.58 (.05)	0.1200	>0.14		54%		54% HK11
CA125	39	649	277	277	971	20	4494	24	24	21	9	84	0.00	<0.001	<0.001	0.933 (.028)	<0.001	>34		60%		94% CA125
HK7 & 10 11	39	4.61	0.50	0.50	15.8	-4.74	82.26	194	-3.84	-3.68	2.07	-10.05	1.07	0.002	<0.001	0.925 (.030)	<0.001	>-1.32		65%		65% HK7 & 10 11
CA125 HK8	39	48.64	15.02	15.02	82.86	-5.66	362.65	186	-5.07	-4.92	1.91	-2.43	0.90	0.0002	<0.001	0.918 (.017)	<0.001	>-2.52		97%		90% CA125 HK8

**We Claim:**

1. A method for detecting a plurality of kallikrein markers associated with ovarian cancer in a patient comprising:
  - 5 (a) obtaining a sample from a patient;
  - (b) detecting in the sample a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
  - (c) comparing the detected amounts with amounts detected for a standard.
- 10 2. A method for diagnosing and monitoring ovarian cancer in a subject comprising detecting in a sample from the subject a plurality of kallikrein markers, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein.
3. A method as claimed in claim 1 or 2 wherein the plurality of kallikrein markers are detected using
  - 15 antibodies that bind to each of the plurality of kallikrein markers or parts thereof
4. A method as claimed in claim 1, 2, 3 which further comprises detecting CA125.
5. A method of detecting ovarian cancer in a patient, the method comprising comparing:
  - (a) levels of a plurality of kallikrein markers, and optionally CA125, in a sample from the
    - 20 patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
    - (b) normal levels of expression of the plurality of kallikrein markers, and optionally CA125, in a control sample, wherein a significant difference in levels of kallikrein markers and optionally CA125, relative to the corresponding normal levels, is indicative of ovarian cancer.
- 25 6. A method for monitoring the progression of ovarian cancer in a patient, the method comprising: (a) detecting in a sample from the patient at a first time point, a plurality of kallikrein markers, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; (b) repeating step (a) at a subsequent point in time; and (c) comparing levels detected in steps (a) and (b), and thereby monitoring the
  - 30 progression of ovarian cancer.
7. A method for determining in a patient whether ovarian cancer has metastasized or is likely to metastasize in the future, the method comprising comparing (a) levels of a plurality of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein
  - 35 10, and kallikrein 11; and (b) normal levels or non-metastatic levels of the kallikrein markers and optionally CA125, in a control sample wherein a significant difference between the levels of expression in the patient sample and the normal levels or non-metastatic levels is an indication that the ovarian cancer has metastasized.
8. A method for assessing the aggressiveness or indolence of ovarian cancer comprising comparing:



- 43 -

- (a) levels of expression of a plurality of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (b) normal levels of expression of the plurality of markers and optionally CA125, in a control sample, wherein a significant difference between the levels in the patient sample and normal levels is an indication that the cancer is aggressive or indolent.
- 5
9. A method for diagnosing and monitoring ovarian cancer in a sample from a subject comprising isolating nucleic acids from the sample; and detecting in the sample polynucleotides encoding a plurality of kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein.
- 10
10. A method as claimed in claim 9 wherein significant differences in the levels of the polynucleotides in the sample compared to a control is indicative of disease, disease stage, and/or prognosis.
11. A method for determining the presence or absence of ovarian cancer in a subject comprising: (a) contacting a sample obtained from the subject with oligonucleotides that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (b) detecting in the sample a level of nucleic acids in the sample that hybridize to the polynucleotides relative to a predetermined cut-off value, and therefrom determining the presence or absence of ovarian cancer in the subject.
- 15
- 20
12. A method as claimed in claim 11, wherein the nucleic acids are mRNA and the levels of nucleic acids are detected by polymerase chain reaction.
13. A method as claimed in claim 11 wherein the nucleic acids are mRNA and the amounts of mRNA are detected using a hybridization technique, employing oligonucleotide probes that hybridize to kallikrein markers, and optionally CA125.
- 25
14. A method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, the method comprising comparing: (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from a patient and exposed to the test agent, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and (b) levels of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient, wherein the sample is not exposed to the test agent, wherein a significant difference in the levels of expression of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the first sample, relative to the second sample, is an indication that the test agent is potentially efficacious for inhibiting ovarian cancer in the patient.
- 30
- 35
15. A method of claim 14 wherein the first and second samples are portions of a single sample obtained from the patient.
16. A method of claim 14 wherein the first and second samples are portions of pooled samples obtained

- 44 -

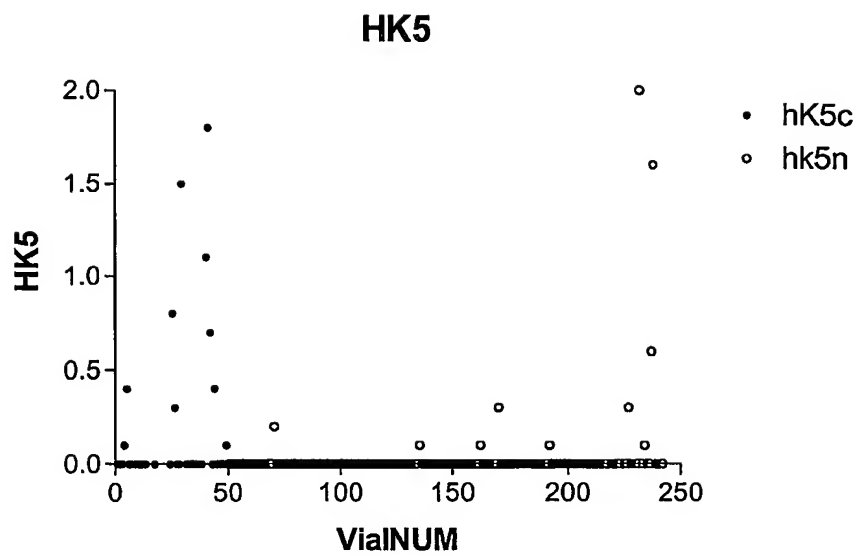
from the patient.

17. A method of assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient, the method comprising comparing: (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from the patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and (b) levels of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient following therapy, wherein a significant difference in the levels of expression of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the second sample, relative to the first sample, is an indication that the therapy is efficacious for inhibiting ovarian cancer in the patient.
18. A method of selecting an agent for inhibiting ovarian cancer in a patient the method comprising (a) obtaining a sample comprising cancer cells from the patient; (b) separately exposing aliquots of the sample in the presence of a plurality of test agents; (c) comparing levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (d) selecting one of the test agents which alters the levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.
19. A method of inhibiting ovarian cancer in a patient, the method comprising (a) obtaining a sample comprising cancer cells from the patient; (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents; (c) comparing levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (d) administering to the patient at least one of the test agents which alters the levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.
20. A method of assessing the ovarian cell carcinogenic potential of a test compound, the method comprising: (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and (b) comparing expression of a plurality of markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and wherein a significant difference in levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot maintained in the presence of the test compound, relative to the aliquot maintained in the absence of the test compound, is an indication that the test compound possesses ovarian cell carcinogenic potential.
21. A method of inhibiting ovarian cancer in a patient at risk for developing ovarian cancer, the method comprising inhibiting expression of genes encoding kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5,

- kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
22. A method of any preceding claim wherein the plurality comprises at least three of the markers.
23. A method of any preceding claim wherein the plurality comprises at least five of the markers.
24. A method of any preceding claim wherein the plurality of kallikrein markers is selected from the  
5 group consisting of kallikrein 5, kallikrein 7, and kallikrein 8; kallikrein 5, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; kallikrein 5, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; or kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.
- 10 25. A method of any proceeding claims wherein the kallikrein markers are kallikrein 7, kallikrein 8, kallikrein 10 and kallirkein 11.
26. A method of any preceding claim wherein the patient sample comprises serum obtained from the patient.
27. A kit for carrying out a method as claimed in any preceding claim.
- 15 28. A kit for assessing whether a patient is afflicted with ovarian cancer, the kit comprising reagents that specifically bind with a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
- 20 29. A kit for assessing the suitability of each of a plurality of agents for inhibiting ovarian cancer in a patient, the kit comprising: (a) the plurality of agents; and (b) reagents for detecting a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
- 25 30. A kit as claimed in claim 28 or 29 wherein the reagents are antibodies that specifically bind with protein or protein fragments corresponding to kallikrein markers and optionally CA125.

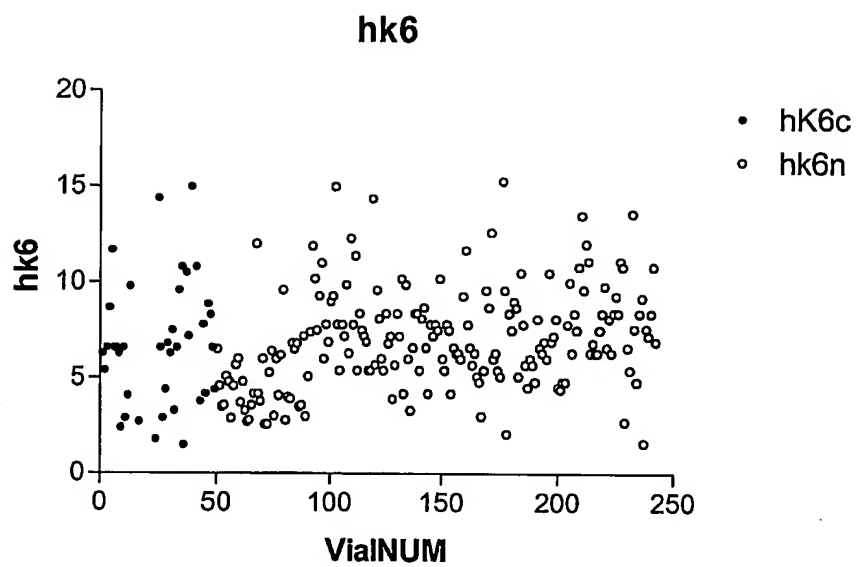
1/8

Figure 1



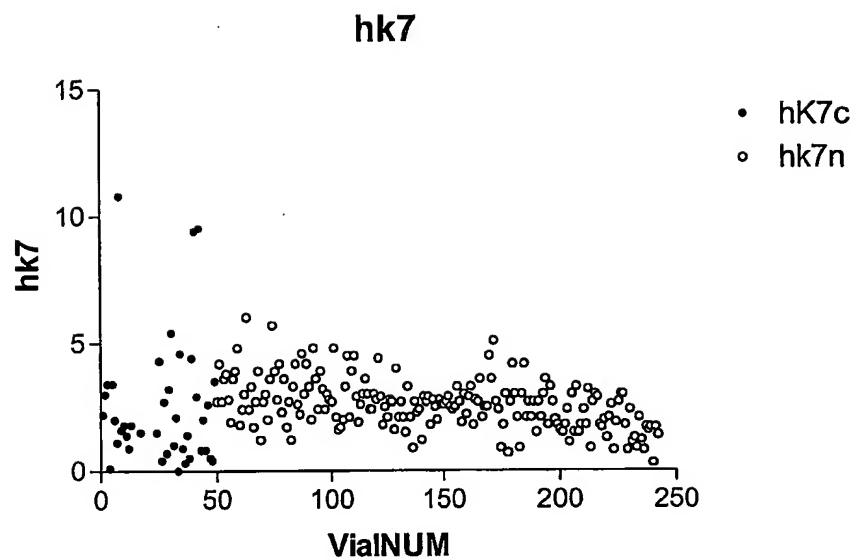
2/8

Figure 2



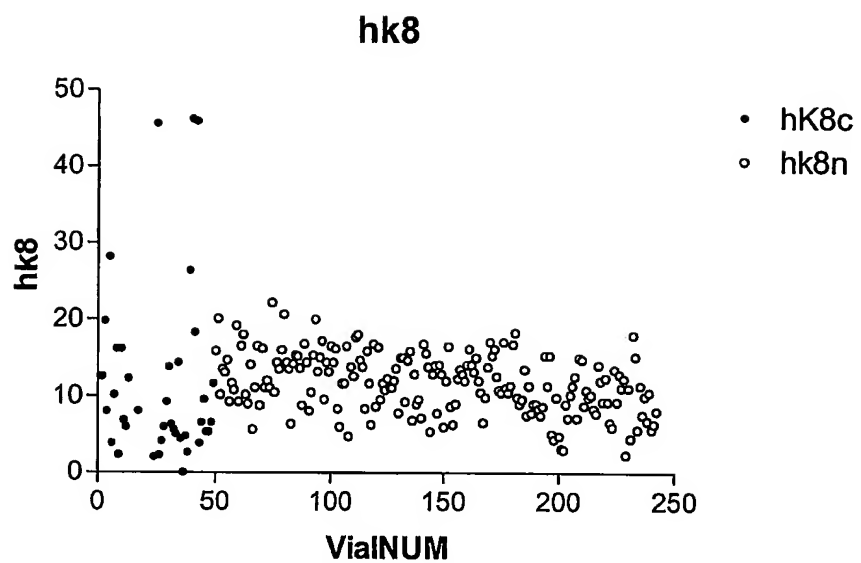
3/8

Figure 3



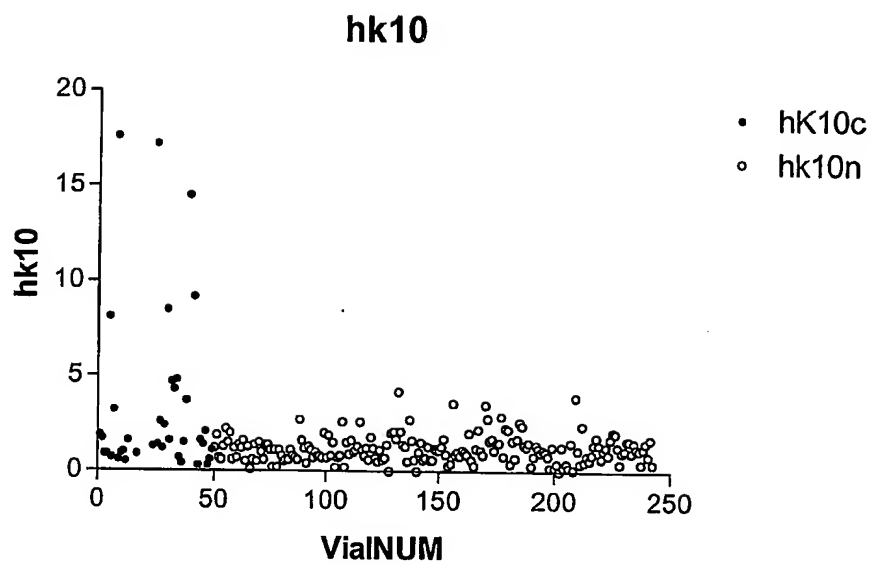
4/8

Figure 4



5/8

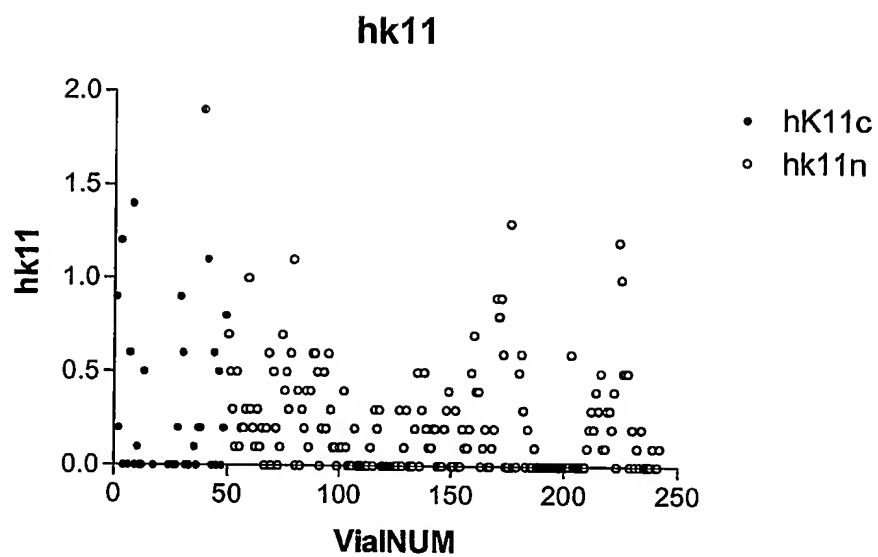
Figure 5





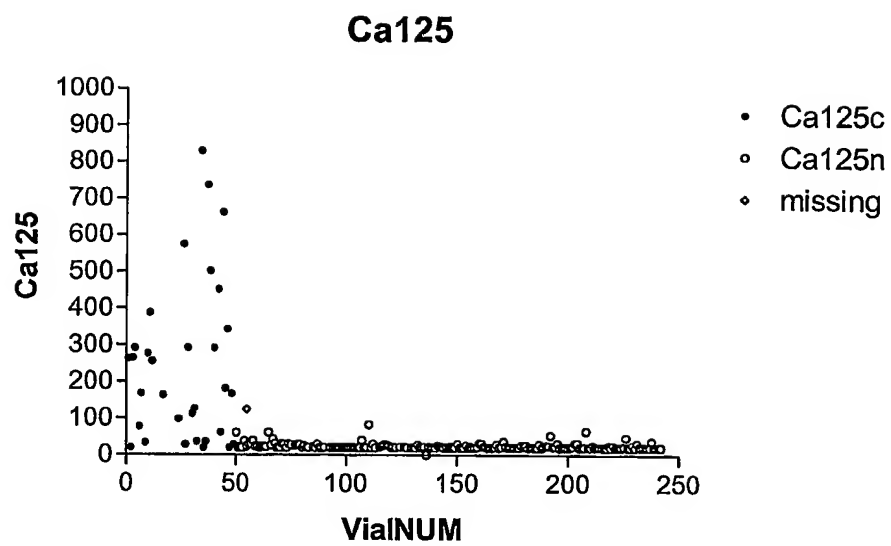
6/8

Figure 6



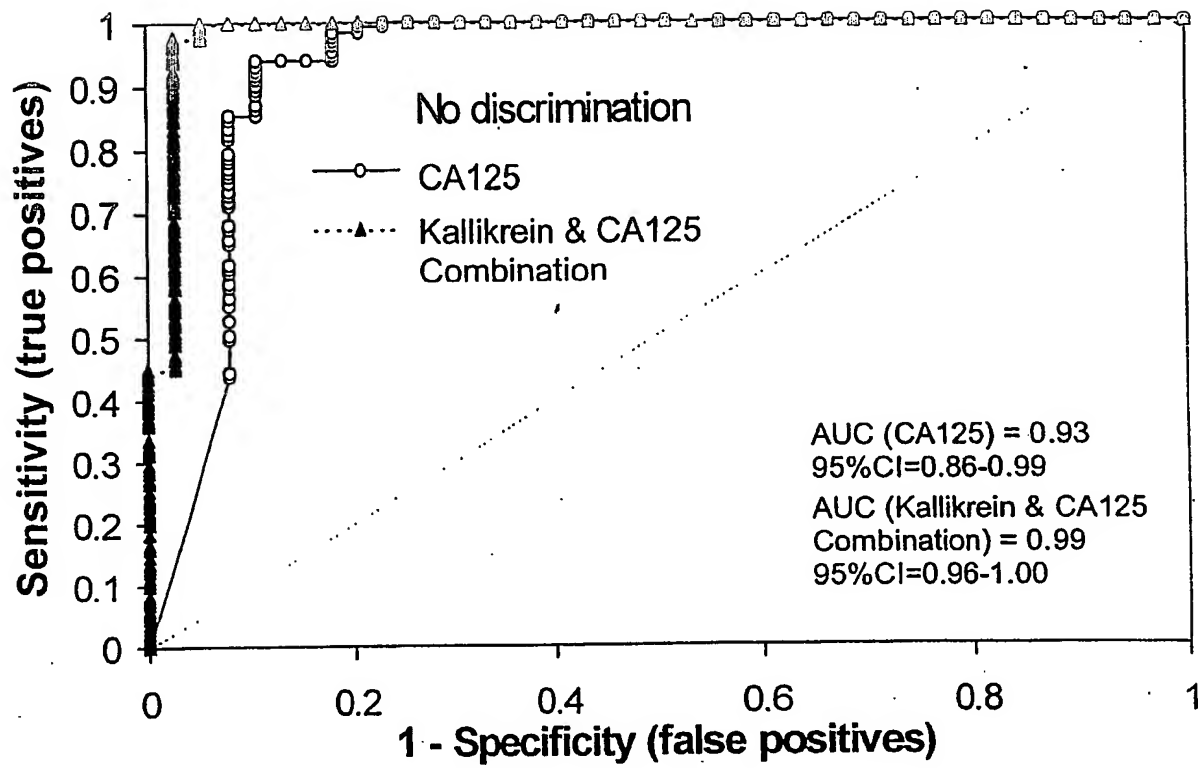
7/8

Figure 7



8/8

Figure 8



## Sequence Listing

SEQ ID NO. 1

CA125 amino acid

```

5      1 mlkpsglpgs ssptrslmtg srstkatpem dsrgltgatls pktstgaivv tehtlpftsp
      61 dktlasptss vvgrrttqslg vmssalpest srgmthseqr tpslspqvn gtpsrmypat
      121 smvsglsspr trtsstegnfn tkeastytlr vettsgpvte kytvptetst tegdstetpw
      181 dtryipvkit spmktfadst askenapvsm tpaettvtds httpgrtnpsf gtlyssfldl
      241 spkgtpnsrg etslelilst tgyffsssep gsaghsrist saplssasv ldnkisetsi
10     301 fsgqsltspl spgvpearas tmpnsaifps mtlснаetса ervrstissl gtpsistkqt
      361 aetiltfbaf aetmdipsth iaktlasewl gspgtlggts tsalttttaps ttlvseentn
      421 hhstsgkete gtlnstmtpl etsapgeese mtatlvpitlg fttldskirs psqvssshpt
      481 relrttgsts grqssstaah gssdilratt sstskasswt sestagqfse pqhtqvwets
      541 psmkterppa stsvaapitt svpsvsgfft tlktsstkgi wleetsadtl igestagptt
15     601 hqfavptgis mtggsestrgs qgtthlltra tassetsadl tlatngvpvs vspavsktaa
      661 gsspppgtkp sytmvasvip etsalqssaf regtslgltp lntrhpfssp epdsaghtki
      721 stsipllssa svledkvsat stfshhkats sittgtpeis tktkpssavl ssmtlснаat
      781 spervrnats plthpspsge etagsvltls tsaettdspn ihptgtltse ssesptlsl
      841 psvsgvkttf sstspthlfl tsgeeteets npsvsgpets vsrvrttlas tsvptpvfpt
20     901 mdtwptrsaq fssshlvsel ratsstsvtn stgsalpkis hltgtatmsq tnrdtfnssa
      961 apgattwpet sprfktglps atttvstsat slsatvmvsk ftspatssme atsirepstt
      1021 ilttettngp gsmavastni pigkgyiteg rldtshlpig ttassetsmd ftmakesvsm
      1081 svpsqsmda agsstpgrts qfvdtfsddv yhltsreiti prdgtssalt pgmtathpps
      1141 pdpgsarstw lgilsspsps ptpkvtmest fstqrvttsm imdtvetsrw nmpnlpstts
25     1201 ltpaniptsg aigkstlvpl dtpspatsle aseggltpls typestntps ihlgahasse
      1261 spstikltma svvkpgsytp ltfpsieith hvstarmays sgsspemtap getntgstwd
      1321 pttiyitttdp kdtssaqvst phsvrtlrtt enhpktesat paaysgspki sspnltspa
      1381 tkawtitdtt ehstqlhytk laekssgfet qsapgpvsavv iptseptigss tleltsdvpg
      1441 eplvlapseq ttitlpmatw lstalteema stldldssps spmatfaifp pmstphels
30     1501 kseadtasair ntdsttldqh lgirslgrtg dlrtvptipl tttwtsvieh stqagdtlsa
      1561 tmsphtvtqs lkdqtsipas aspshlvey pelgtqgrss seattfwkps tdtlsreiet
      1621 gptniqstpp mdntttgsss sgvtlgiahl pigtsspaet stnmalerrs statvsmagt
      1681 mglvtsapq rsisqslgrv ssvlsestte gvtddskgss prlntqgnta lsslepsy
      1741 egsgmstsis ltasspttdv efiggstfwt kevttvmtsd iskssartes ssatlmstal
35     1801 gstentgkek lrtasmdlps ptpsmevtpw isltlsnapn ttdsldlshg vhtssagtl
      1861 tdrslntgvt rasrlengsd tsakslsmgn sthtsmtdte ksevsstsih rpetvsgae
      1921 ttltstpgnr aistltpfss ipveevistg itsgpdinsa pmthspitpp tivwtstgti
      1981 eqstqplhav ssekvsvtqg stpyvnsavv saspthensv ssgsstsspy ssaslealds
      2041 tisrrnais wldtlttelp ttwtpstsls ealsaghsqv snpsstttf plfsaasta
40     2101 akqrnpetet hgpqmtaast lntdassvtg lsetpvgsasi ssevpplmai tarsdvsglt
      2161 sestansplg tassagtklt rtislptses lvsfrmkdp wtvsiplgsh pttntetsip
      2221 vnsagppgls tvasdvidtp sdgaesiptv sfspspdtev ttishfpekt thsfrtissl
      2281 theltsrvtp ipgdwmssam stkptgasp itlgerrtit saapttspiv ltasftetst
      2341 vsldnettvk tsdildarkn nelpsdssss sdintiesias stmdvktas isptsisgmt
45     2401 asspslfs drpqvptstt etntatpsv ssntysldgg snvggtptsl ppftithpve
      2461 tssallawsr pvrtfstmvs tdtasgenpt ssnsvvtsvp apgtwasvgs ttdlpamgfl
      2521 ktspageahs llastiepat aftphlsaav vtgssatsea sltttseka ihaspgtptt
      2581 ptsganwets atpesllvvt etsdtltlsk ilvtdtilfs tvstppskfp stgtlsgasf
      2641 ptllpdtpai pltateptss latsfdstpl vtiasdslgt vpettltmse tngdalvfk
50     2701 tvsnpdrrsip gitigqvtes plhpsstsp kivaprntty egsitvalst lpagttgslv
      2761 fsqssenset talvdssagl erasvmltt gsqgmassgg irsgsthtg tktfsslplt
      2821 mnpgevtams eitnrltat qstapkgipv kptsaesgll tpvsassps kafaslttap
      2881 pstwgipqst ltfevsevp ldtksaslt pgqslntipd sdastassel skspeknpra
      2941 xmtstskais assfgstgft etpegsasps magheprvpt sgtgdpryas eemsypdpks
55     3001 assamtstsl asklttlfst gqaarsgss spislsteke tsflseptast srktslflgp
      3061 smarqpnilv hlqtsaltls ptstlnmsqe eppeltssqt iaeeegttae tqtlfttpe
      3121 tptslpvs pteptarrks spetwassis vpaktslvet tdtglvttik mssqaaggs
      3181 twpapaeetg tspagtepps pevsttlkim sskepsispe irstvrnsqw ktpettvpme
      3241 ttvepvtlqs talgsgtsi shlptgttsp tksptenmla tervslspsp peawtnlysg

```

	3301	tpggtrqsla	tmssvslesp	tarsitgtgq	qsspelvskt	tgmefsmwhg	stgggttgdh
	3361	vsletssnil	edpvtspnsv	ssltckskhk	tetwvsttai	pstvlmnmk	aaeqqtzsv
	3421	deaysstssw	sdqtsqsdit	lgaspdvnt	lyitstaqt	slvslpsgdq	gitsltnpsg
	3481	gktsasssvt	speigletlr	anvsavksdi	aptaghlst	sspaevsild	vttaptpgis
5	3541	ttittmgtns	istttpnpev	gmstmdstpa	terrttsteh	pstwsstaas	dswtvtmdts
	3601	nlkvarspgt	istmhttsfl	assteldsm	tphgritvig	tslvtppsda	savktetsts
	3661	ertlspsttt	astpietfss	vqrmsisvpd	ilstswtpss	teaedvpvsm	ntvdhastkt
	3721	dpntplstfl	fdslstldwd	tgrslssata	ttsapqgatt	pgeltletmi	spatsqlpfs
	3781	ighitsavtp	aamarssgv	fsspdptskk	aeqtstqlpt	ttsahpggvp	rsaattldvi
10	3841	phtaktptat	fqrqggtalt	tearatsdsw	nekekstpsa	pwitemmns	sedtikevts
	3901	sssvlkdpay	aghlklgiwdd	fipkfgkaah	mrelpllspp	qdkaihpst	ntvettgwvt
	3961	ssehashsti	pahsasklt	spvvtststre	qaivsmsttt	wpestrarte	pnsfltielr
	4021	dvspymdtss	ttqtsiissp	gstaitkgpr	teitsskris	ssflaqsmrs	sdspseaitr
	4081	lsnfpmates	ggmilamqts	ppgatslsap	tldtsatasw	tgtplattq	ftysekttlf
15	4141	skgpedtsqp	spsveetss	ssslvpihat	tspsnillts	qghspsstpp	vtsvflsets
	4201	glgkttmdsr	islepgtslp	pnlsststagea	lstyeasrdt	kaihhsadta	vtmneatsse
	4261	yspipgthkp	skatsplvts	himgditsst	svfgssette	ietvssvngq	lqerstsqva
	4321	ssatetstvi	thvssgdatt	hvtktqatfs	sgtsissphq	fitstntftd	vstnpstali
	4381	mtessgvtit	tqgtgtgaat	qgpylldtst	mpyltetpla	vtpdfmqsek	ttliskgpkd
20	4441	vtwtspssva	etsypsltp	flvtittippt	stlqgqhtss	pvsatsvlt	glvkttdmln
	4501	tsmepvttnsp	qnlmnpnei	latlaattddi	etihpsinka	vtmngtassa	hvlhstlpvs
	4561	sepstatatpm	vpassmgdal	asisipgset	tdiegeptss	ltagrkenat	lqemnsttes
	4621	niilsnsvsg	aiteatkmev	psfdatfipt	paqstkfpdi	fsvassrlsn	spmtisthm
	4681	tttqtgssga	tskiplaldt	stletsagtp	svvttegfahs	kittamndv	kdvsgttnppf
25	4741	qdeasspsag	apvlvtllps	svaftpqwhs	tspsvsmssv	ltsslvktag	kvdtsletvt
	4801	sspqmsntl	ddisvtsaat	tdietthpsi	ntvvtnvggt	gsafeshstv	saypepskvt
	4861	spvttstme	dttsirsipk	skttrtete	ttssltptklr	etsisqeits	stetstvyk
	4921	eltgattevs	rtadvtsst	sfpqpdqstv	sldistetnt	rlstspimte	saeittittqt
	4981	gphgatsqdt	tmdpsnttp	qagihsamth	gfsqldvttl	msripqdvsw	tspssvdkts
30	5041	spssflsapa	mttspssist	lpedklsspm	tslltsglvk	itdlrlrle	pvtsslpnfs
	5101	stsdkilats	kdskdtkelf	psinteetnv	kannsgghesh	spaladsetp	kattgmvtit
	5161	tvqdpapsts	mpvhgssett	nikreptyfl	tprlretsts	qessfptdts	fllskvptgt
	5221	itevsstgvn	ssskistpdh	dkstvppdtf	tgeiprvfts	siktksaemt	ittqasppes
	5281	ashstlpldt	sttlsgqgth	stvtqgfpys	evttlmgmpp	gnvswmttpp	veetssvss
35	5341	msspamtaps	pvsstspqsi	psplpvtal	ptsvlvttd	vlgttspeav	tsppnlssi
	5401	therpatyk	tahteamhh	stntavtnvg	tsqsgghksqs	svladsetsk	atplmstst
	5461	lgdtsvstst	pnlsqtnqiq	teptaslspr	lresstsekt	ssttetntaf	syptgaitq
	5521	asrteissrr	tsisdldrpt	iapdistgmi	trlftspimt	ksaemtvtq	tttppgateqg
	5581	ilpwtsttl	fqqgthstvs	qgfhphseitt	lrsrtpgdvs	wmttppveet	ssgflmsps
40	5641	mtspspvst	spesipsspl	pvtalltsvl	vttnvlgtt	spetvtspp	nlsaptqerl
	5701	tttkdtahte	amhasmhtnt	avanvgtsis	ghesqssvpa	dshtskatsp	mgitfamgdt
	5761	svststpaff	etriqtstss	slipglrdtr	tseeintvte	tsvlsevp	tttvevrte
	5821	vtssrttis	gpdhskmspy	istetitrls	tfpfvtgste	maitnqtgpi	gtisqatl
	5881	dtssstasweg	thspvtgrfp	hseetttsmr	stkgvswqsp	psveetssps	spvplpait
45	5941	hsslysavsg	ssptsalpvt	sltsgrrrkt	idmldthsel	vtsslpssas	fsgelstsea
	6001	stntetihfs	entaetnmgt	tnsmhklhss	vsihsqpsgh	tppkvtgsmm	edaivststp
	6061	gspektknvdr	dstspltpel	kedstaltvm	sttesntvfs	svsldaatev	sraevtyydp
	6121	tfmpasaqst	kspdispeas	sshsnspllt	isthktiatq	tgpsgvtslg	qltldtstia
	6181	tsagtpstart	qdfvdsetts	vmndldndvl	ktspfsaeaa	nsllsqapll	vttapsvpts
50	6241	tlqehstss	svstsvptpt	lakitdmtdn	lepvtvspqn	lrlntlatsea	tttdthmhp
	6301	intamanvgt	tsppnefyft	vspdsdpyka	tsavvitsts	gdsivstsm	rsamakkies
	6361	ettflslifrl	retstsqkig	sssdststvf	kaftaattev	steltsssr	tslqgtekpt
	6421	mtpdtstsv	tmlstfaglt	kseertiatq	tgphratsqg	tltwdtsitt	sqagthsam
	6481	hgfsqldlsl	ltsrvpeyis	gtsppsvekt	ssssllslp	aitspspvpt	tlpesrpss
55	6541	vhltslptsg	lvkttdmlas	vaslppnlgs	tskhkpttse	dikdtekmy	stniavtnvg
	6601	tttsekesys	svpayseppk	vtspmvtsfn	irdtivistm	pgsseitrie	mestfsvahg
	6661	lkgtsstsqdp	ivsteksav	hklttgatet	stevassrr	tsipgpdhat	espdistevi
	6721	pslpislgit	essnmittir	tgpplgstsq	gtftldtptt	ssragthsm	tqefphsemt
	6781	tvnmkdpell	swtippsie	tsfssslmps	pamtsppvss	tlpktihttp	spmtalltps
60	6841	lvmttdtlgt	spepttsspp	nlststshvil	ttddedtaie	amhpststaa	tnvettcsgh
	6901	gsqssvltlds	ektkatapmd	ttstmghtt	stsmssvset	tkikrestys	ltplrltsi

5 6961 sqnasfsttd sivilsevpptg ttaevsrtev tssgrtsipg psqstvlpei strtmtrlfa  
 7021 sptmtesaem tiptqtgpgg stsqdtltld tatkksqakt hstltqrfph semttlmsrg  
 7081 pgdmswqssp slenpsslps llslpattsp ppissltlptv isssplpvts lltsspvttt  
 7141 dmlhtspelv tssppklsht sderlttgkd ttnteavhps tntaasnvei psfghespps  
 7201 aladsetska tpmfittstq edttvaistp hfletsriqk esisslsplk retgssvets  
 7261 saietsavls evsigattei srtevtsssr tsisgsaest mlpeisttrk ikkfptspil  
 7321 aessemtikt qtsppgstse stfldtstt pslvithstm tqrlphseit tlvsrgagdv  
 7381 prpsslpvee tppssqsls samispspvs stlpasshss sasvtspltp gqvkttevl  
 7441 asaepetssp pslsstsvei latsevttdt ekihpfpnta vtkvgtsess hespsavlpd  
 10 7501 settkatsam gtisimgdts vstltplasn trkqiqsepas slttrlrts tseetslate  
 7561 antvlskvst gattevsrte aifsrtssms gpegstmsqd isigtipris assvltesak  
 7621 mtittqtgpps estlestlnl ntattpswve thsiviqgfp hpemttmgr gpggvswpsp  
 7681 pfvketspps splslpavts phpvsttfla hippslpvt sltsgpatt tdilgtstep  
 7741 gtsassslst tsherltyk dtahteavhp stntggtnva tssgyksgs svladsppmc  
 15 7801 ttstmgdtsv ltstpafler rriqtelass ltpglressg segtssgkkm stvlsvkvtg  
 7861 atteiskedv tsipppagst ispdistrtv swfstspvmt esaeitmnts tsplgattqg  
 7921 tstlatsott sltmthstis qgfshsqmst lmrrgpedvs wmsppllekt rpsfslmssp  
 7981 attspspvss tlpesissss lpvtsalltg lakttdmlhk ssepvtspna nlsstsvail  
 8041 atsevttdte kthpsnrtv tdvgtssaghs estsfvlads qtskvtspmv itstmedtsv  
 20 8101 ststpgffet sriqteptss ltlglrktss segtslatem stvlsgvptg ataevsrtev  
 8161 tssrtsisg faqltvspet stetitrlpt ssimtesaem miktqtdppg stpesthtvd  
 8221 isttpnwvet hstvtqrfsh semttlvers pgdmlwpsqs aveetssass llslpattsp  
 8281 spvstlved fpsaslpvts lltpglvitt drmgisrepg tsstsnlss sherlttled  
 8341 tvdtedmpps thtavnrtv sisghesqss vlsdsetpka tspmgttytm getsvissts  
 25 8401 dffetsriqi eptssltsgl retsseris sategstvlsv evpsgattev xrevisarg  
 8461 tsmgpdqft ispdisteai trlstspimt esaesaitie tgspgatseg tltldtstt  
 8521 fwsgthstas pgfshsemt lmsrtpgdvp wpslpveea ssvssslsp amtstsfssa  
 8581 lpesissph pvtalltlgp vkttdmlrts sepetssppn lsssaaila tsevtkdrek  
 8641 ihpsantpvr nvgtvlykhl spssvladv tktptspmat tstlgtsvs sttpafpmt  
 30 8701 mtqptssls glreistsqe tssatersas lsgmptgatt kvsrtealsl grtstpgpaq  
 8761 stispeiste titristplt ttgsaemtit pktghagass qgtftldtss raswpgthsa  
 8821 athrsphagm ttpmsrgped vswpsrpsve ktsspslsvs lsavtspsp ystpsesshs  
 8881 splrvtslft pvmmkttdml dtselepvtss ppsmnitsde slatskatne teaiqlsent  
 8941 avtqmgtsa rgefyspyg lpepskvts vvtssstikdi vsttipasse itriemests  
 35 9001 tltptprets tsqeihsatk pstvpykalt satiedsmtd vmsssrsgsp dqtmsqdis  
 9061 sevitrlsts pikaestemt ittqtgspga tsrgtltldt sttfmsgths taaggfshsq  
 9121 mtalmsrtpp dvpwlsghpsv eeassasfsl sspvmtsssp vssltpldsih ssplpvtll  
 9181 tsglvkttel lgtsssepets sppnlsssta eilattevtt dteklemtv vtsgythesp  
 9241 ssvladvtt katssmgity ptgdnvlt tpaftsdsri qtkskslst glmetsisee  
 40 9301 tssatekstv lssvptgatt evsrteaiss srtsipppaq stmsstdtame titristplt  
 9361 rkestdmait pktgpgats qgtftldess taswpgthsa ttqrfpqsvv ttpmsrgped  
 9421 vswpsplsve knsppsslv ssvtspsp ystpsgsshs spvpvtslft slmmkatdml  
 9481 daslepetts apnmnitsde slatskatte teaihvfent aashvettsa teelyssspg  
 9541 fseptkvisp vvtsssiridn mvsttmgss gitrieiesm sstlpglret rtsqdisst  
 45 9601 etstvlykms sgatpevsrt evmpssrtsi ppgaqstmsl disdevvtrl stspimtesa  
 9661 eititttqgy slatsqvltl lgtsmftlsg thstmsqgls hsemtnlmar gpeslwtsp  
 9721 rfvettrss sltslpltt lsvsstlld sspsslpvt slilpglvkt tevltdasap  
 9781 ktsspnls tsveipatse imtdtekihp ssntavakvr tsssvhshs svladsetti  
 9841 tipsmgitsa vddttvftsn pafsettripp teptfsltp fretstseet tsitetsavl  
 50 9901 ygvptsatte vsmteimssn rthipdsdqs tmspditev itrssssmm sestqmtitt  
 9961 qksspgataq stltlattta plarthstvp prflhsemt lmsrpenps wksspfvekt  
 10021 sssslslslp vttspsvsst lpqsspsf avtsalltpgm vkttdtstep gtslspnlsg  
 10081 tsveilaase vttdekihp sssmavtnvg tssghelys svaihsepsk atypvgtpss  
 10141 maetsistem panfettgfe aepfshltsg frktnmsldt ssvtptntps spgsthllgs  
 55 10201 sktdftssak tssdpwppas qyteipvdii tpfnaspsit estgitsfpe srtfmsvtes  
 10261 thhlstdllp saetistgtv mpsslseamt fattgvprai sgsgspfsrt esgpgdatls  
 10321 tiaeslpst pvpfssstft ttdsstipal heitsssatp yrvdtslgt ssttegrlvm  
 10381 vstldtssqp grtsstpild trmtesvelg tvtsayqvps lstrlrltdg imehickipn  
 10441 eaahrgttrp vkpggtstap aspkglltgg tkrmetttta lktttalkt tsratltsv  
 60 10501 ytptlgtltp lnasrqmast iltemmittp yvfpdvpett sslatslgae tatalprtt  
 10561 svlnresett aslvrsrgae rsvpiqtlav sssepdtas wvhaetip tvskttpnff

10621 hseltdvssst atshgadvs aiptnispse ldaltplvti sgttdtsttfp tltksphete  
 10681 trttwlthpa etsstiprti pnfshhesda tpsiatspga etssaipimt vspgaedlvt  
 10741 sqvtssgtdr nmtiptltls pgepktiasl vthpeagtss aiptstispa vsrlvtsmvt  
 10801 slaaktsttn raltnsppgep attvslvthp aqtsptvpwt tsiffhsksd ttpsmtshg  
 5 10861 aesssavptp tvstevpgvv tplvtssrav isttipiltl spgepettps matshgeaas  
 10921 saiptptvsp gvpvvtslv tssravtatt ipiltfslge pettpsmats hgteagsavp  
 10981 tvlpevpvpmv tsavssrav tstitlptltl spgepettps matshgaeas stvptvapev  
 11041 pgvvtslvts ssgvnatsip tlllspgele ttpsmatshg aeassavptp tvspgvsgrv  
 11101 tplvtssrav tstitipiltl ssepettps matshgveas savltvapev pgmvtslvt  
 10 11161 sravtsttip tltissdepe tttslvthse akmisaiptl avspvtqglv tslytssgse  
 11221 tsafsnltva ssgpetidsw vahpgteass vvpiltvtstg epftnlslvt hpaessstlp  
 11281 rttstrfshse ldtmptvts peaeassais ttispgipgv ltslytssgr disatfptvp  
 11341 espheaseata swvthpavts ttvprttvny shsepdtts iatpagaat sdfptitvsp  
 11401 dvpdmvtsqv tssgtdtsit iptltlssge petttsfity sethtssaip tlpvsggask  
 15 11461 mltalvisag tdstttfptl tetpyeett aiqlhpaet ntmvpkttk fshksdttl  
 11521 pvaltspgpe assavsttti spdmsdlvts lvpssgtdts ttfptlsetp yepettvtwl  
 11581 thpaetsttv sgtipnfshr gsdtpasmvt spgvdrsgv ptttippsip gvvtsqvtss  
 11641 atdtstaip ltpspgepet tassathpgt qtgftvpipt vpssepdtma swvthppqts  
 20 11701 tpvrttssf shsspatpv matsprteas savlttispg apemvtsqit ssgaatsttv  
 11761 ptlthspgmp ettallsthp rtgtsktfpa stvfpgvset tasltirpga etstaltptqt  
 11821 tsslftllvt gtsrvdlst aspvsakta plsthpgtet stmiptstls lglltettgll  
 11881 atsssaetst stltltvspa vglssasit tdkpqtvtsw ntetspsvts vgppefartv  
 11941 tgtmtlips emptppktsh gegvspttil rttmveatnl attgsseptva kttttfntla  
 12001 gslftpltt gmstlaseev tsrtsynhrs wisttsynr rywtpatstp tdstfpgis  
 25 12061 tssipsstaa tvpfmvpftl nftitnlqye edmrhpgsrk fnaterelqg llkplfrnss  
 12121 leylysgcrl aslrpekds amavdaicth rpdpedlgld rerlywelsn ltngigelgp  
 12181 ytlrnslyv ngfthrsmp tdstpgtstv dvgtsgtpps speptaagpl lmpftlnfti  
 12241 tnlqeedmr rtgsrkfntm esvlqgllkp lfkntsvgpl ysgcrltlr pekdgatgv  
 30 12301 daicthrlp kspglnreq ywelsklnd ieelgpytld rnslyvngft hqssvstst  
 12361 pgtstvdlt sgtpsalssp timaagpllv pftlnftitn lqygdmghp gsrkfntter  
 12421 vlqgllgpif kntsvgplys gcrllsrse kdgaatgvda icihhldpks pglrnrlyw  
 12481 elsqtlngik elgpytldrn slyvngfthr tsvpttstpg tstvdltsg tpfslpspat  
 12541 agpllvftl nftitnlkye edmrhpgsrk fnttervlt lpgpmfknts vgllysgcrl  
 35 12601 tllrsekda atgvdaicth rldpkspgld reqlywelsq ltngikelgp ytlrnslyv  
 12661 ngfthwipvp tsstpgtstv dlsgtpps psptaagpl vpftlnftit nlqeedmhh  
 12721 pgsrkfntte rvlqgllgpm fkntsvgply sgcrlllrs ekdgaatgv aicthrlpdk  
 12781 spgvdrqly welsqltngi kelgpytldr nslyvngfth qtsapntstp gtsvdlgts  
 12841 gtpsslpst sagpllvft lntitnlqy eedmrhpgsr kfnttervlt gllkplfkst  
 12901 svgplysgcr ltlrsekdg aatgvdaic hrlpdkspgv dreqlywels qltngikelg  
 40 12961 pytlrnsly vngfthqtsa pntstpgtst vdlgtsgtps slpsptsagg llvpftlnft  
 13021 itnlqeedm hhpgrkntf tervlqgllg pmfkntsvgl lysgcrltl rpekngaag  
 13081 mdaicshrl pkspglnreq lywelsqlth gikelgpytl drnslyvngf thrssvapt  
 13141 tpgtstvdlt sgtpsslp pttavpllv ftlnftitnl qygdmrhpgr srkfntterv  
 13201 lqgllgplfk nssvplyeg crllsrsek dgaatgvda cthhlnpqp gldreqlwq  
 45 13261 lqmtngike lqpytldrn lyvngfthrs sglttstpw stvdlgtsgt pspvpsptta  
 13321 gppllvftln ftitnlqye dmhrpgsrkf nttervltg lspifknssv gplysgcrlt  
 13381 slrpekdgaa tgmavclyh pnpkrpgldr eqlywelsq thnitelgpy slrdsalyv  
 13441 gfthqnsvpt ttpgtstvy wattgtpsf pghtepgpl ipftfnftit nlhyeenmqh  
 50 13501 pgsrkfntte rvlqgllkpl fkntsvgply sgcrllslrp ekdgaatgmd avclyhpnk  
 13561 rpgldreql welsqlthni telgpytldr dslyvngfth qnsvptstp gtsvwyatt  
 13621 gtpssfpght epgpilipft nftitnlhy eenmqhpgsr kfnttervlt gllkplfknt  
 13681 svgplysgcr ltlrpekhe aatgvdtict hrvdipgpl drerlywels qltnsitelg  
 13741 pytlrdsly vngfnprssv pttstpgtst vhlatsgtps slpghtapvp llipftlnft  
 55 13801 itnlhyeenm qhpgsrkfnt tervlqgllk plfkntsvgp lysgcrltl rpekheaag  
 13861 vdticthrvd pigpplxex lywelsxlt xixelgpytl drxslyvngf thxaxptts  
 13921 tpgtstvxg tsptsspxp xtsagpllv ftlnftitnl qyeedmhhpg srkfntterv  
 13981 lqgllgpmfk ntsvglyeg crllrpek ngaatgm dai cshrlpdkp gldreqlw  
 14041 lsqthgike lqpytldrn lyvngfthrs svaptstpgt stvdlgtsgt psslpstpta  
 60 14101 vpllvftln ftitnlqye dmhrpgsrkf nttervltg lqplfkns vplysgcrl  
 14161 slrsekdgaa tgvdaicthh lnpqspgldr eqlywelsqm tngikelgpy tldrnslyv  
 14221 gfthrsaglt ttpwtstvd lgtsgtspv pspttagpl vpftlnftit nlqeedmhr

14281 pgsrkfnate rvlqgllspi fknssvvgply sgcrltslrp ekdgaatgmd avclyhpnkp  
 14341 rpgldreqly welsqlthni telgpysldr dslyvngfth qssmtttrtp dtstmlhats  
 14401 rtpaslsqpt tasp1lvlt inctitnlqy eedmrtrgsr kfntmesvlq gllkplfknt  
 14461 svqplysgcr ltllrpkkdg aatgvdaict hrldpkspgl nreqlywels klndieelg  
 5 14521 pytlrnsly vngfthqssv sttstpgtst vdlrtsgtps slssptimxx xp1lxpftxn  
 14581 xtitalnxxxx xmxpgsrkf nttervlqgl lrplfkntsv sslysgcrlt llrpekdgaa  
 14641 trvdaactyr pdpkspgl dr eqlywelsql thsitelgpy tldrsvlyvn gfnprssvpt  
 14701 tstpgtstvh latsgtpssl pgghtxxxpl xpftxnxtit nlxxxxmxx pgarkfntte  
 14761 rvlqgllkpl fnssleyly sgcrilaslrp ekdssamavd aicthrpdp d1gldrerly  
 10 14821 welsnltnqi qelgpytl dr nslyvngfth xssglttstp wstsvdlgts gtpspvpspt  
 14881 tagpllvpt lnftitnlqy eedmhrggsr rfnttervlq glltplfknt svqplysgcr  
 14941 ltllrpekqe aatgvdtict hrldp1pgpl drerlywels qltnsitelg pytlrdrslly  
 15001 vngfnpwssv pttstpgtst vhlatsgtps slpghtapvp llipftlnft itdlhyeenm  
 15061 qhpgsrkfnt tervlqgllk plfkstsvgp lysgcrltll rpekghaatg vdaictlrl d  
 15 15121 ptgpgldrer lywelsqltn svtelgpytl drdslyvngf thrssvptts ipgtsavhle  
 15181 tsqtpas1pg htapggllvp ftlnftitnl qyeedmrhpg srkfstterv lqgllkplfk  
 15241 ntavsslysg crltllrpek dgaatrvdav cthrpdpksp gllderlywk lsq1thgite  
 15301 lgpytlrhrs lyvngfthqs smttrtpdt stmh1atsrt paslsqptta spllvltin  
 15361 ftitnlryee nmhpgsrkf nttervlqgl lrpvfkntsv gplysgcrlt tlrpkkdga  
 20 15421 tkvdaictyr pdpkspgl dr eqlywelsql thsitelgpy tqdrdslyvn gthrssvpt  
 15481 tsipgtsavh letsqtpasl pghtapggll vpftlnftit nlqyeedmrh pgarkfntte  
 15541 rvlqgllkpl fkstsvgply sgcr1tllrp ekrgaatgvd ticthrl dpl npgldreqly  
 15601 welskltrgi ielgpyll dr gslyvngfth rtsvpttstp gtstvd1gts gtpfslpspa  
 25 15661 xxxpllxpft xnxtitnlxx xmxmxxpgsr kfnttervlq tllgpmfknt svgllysgcr  
 15721 ltllrsekdg aatgvdaict hrldpkspgv dreqlywels qltnqike1g pytlrdrslly  
 15781 vngfthwipv pttstpgtst vdlsgtpps lpspttagpl lvpftlnfti tnlkyeedmh  
 15841 cpgsrkfntt ervlqsl1gp mfkntsvgpl ysgcrltllr sekdgatgv daicthrl d  
 15901 kspgvdreql ywelsqltnq ikelgpytl dr nslyvngfth hqtsapntst pgtstvd1gt  
 15961 sgtps1psp txxxpllxpf txnxtitnlx xxxmxxpgs rkfnttexvl qgllxpxfk  
 30 16021 xsvgxlysgc rltxlrxekx gaatgxdaic xhxxpxkxpg lxxexlywel s1ltxxixel  
 16081 gpytlr1xsl yvngfthwip vptstpgts tvdlsgtpps slpspttagp llvpftlnft  
 16141 itnlkyeedm hcpgsrkfnt tervlqsl1g pmfkntsvgp lysgcrlt1s rsekdgatg  
 16201 vdaicthrvd pkspgvdr eq lywelsqltn gikelgpytl drnslyvngf thqtsapnts  
 16261 tpgtstvx1g tsqtpssxpx xtsagpllv ftlnftitnl qyeedmhpg srkfntterv  
 35 16321 lqgllgpmfk ntavgllysg crltllrpek ngattgm dai cthrl d ksp glxxexlywe  
 16381 lsx1ltxxix lgytlr1xsl lyvngfthxx xpttstpgt stvxxtstg psspxxtxx  
 16441 xp1lxpftxn xtitalnxxxx xmxpgsrkf nttervlqgl lkplfrnssl eylysgcrla  
 16501 slrpekdsaa mavdaicthr dpdldgl dr erlywelsnl tngiqelgpy tldrdslyvn  
 16561 gfthrrsmpt tstpgtstvd vgtsgtpss pspttagpl ipftlnftit nlqygedmgh  
 40 16621 pgsrkfntte rvlqgllgpi fkntsvgply sgcr1tslrs ekdgaatgv d aic1h1ldpk  
 16681 spglnrerly welsqltnqi kelgpytl dr nslyvngfth rtsvpttstp gtstvd1gts  
 16741 gtpfslpspa tagpllvlt lnftitnlky eedmhrggsr kfnttervlq tllgpmfknt  
 16801 svgllysgcr ltllrsekdg aatgvdaict hrldpkspgl xxexlywels xltxxixelg  
 16861 pytlr1xsl yvngfthxxsx pttstpgtst vxxtstgtps sxpxxtxxx llxpftxnx  
 45 16921 itnlxxxxm xxpgsrkfnt tervlqgllr pvfkntsvgp lysgcrltll rpkkdgaatk  
 16981 vdaic1xrp d pkspgl dr eq lywelsqltn sitelgpytq drdslyvngf thrssvptts  
 17041 ipgtsavhle ttgtpssfpg htepgpllip ftfnftitnl ryeenmhpg srkfntterv  
 17101 lqglltplfk ntavgllysg crltllrpek qeaatgvdti cthrvdp1g gllderlywe  
 17161 lsq1tnsite lgpytlrdr ds lyvdgfnpws svpttstpgt stvhlatsgt p1plpghtap  
 50 17221 vp1lipftln ftitdlhyee nmqhpgrkf nttervlqgl lkplfkstsv gplysgcrlt  
 17281 llrpekghaa tgvdaictlr ldtpggldr erlywelsql tnsitelgpy tldrdslyvn  
 17341 gfnpwssvpt tstpgtstvh latsgtpssl pghttagpl vpftlnftit nlkyeedmhc  
 17401 pgsrkfntte rvlqslh1gm fkntsvgply sgcr1tllrs ekdgaatgv d aicthrl d  
 17461 spglxxexly welsx1ltxxi xelgpytl dr xslyvngfth xxsxpttstp gtstvxxts  
 55 17521 gtpsspxxt xxxpllxpft xnxtitnlxx xmxmxxpgsr kfnttexvl gllxpxfkx  
 17581 svgxlysgcr ltxlrxekx aatgxdaicx hxxpxkxpgl xxexlywels xltnsitelg  
 17641 pytlrdrslly vngfthrrsm pttstpgtsa vhlatsgtpa slpghtapgp llvpftlnft  
 17701 itnlqyeedm rhpgsrkfnt tervlqgllk plfkstsvgp lysgcrltll rpekrgaatg  
 17761 vdticthrl d plnpglxxex lywelsxltx xixelgpytl drxslyvngf thxxsxp1ts  
 60 17821 tpgtstvx1g tsqtpssxpx txxxpllxp ftmxtitnl xxxmxxpgs rkfnttexv  
 17881 lqgllxpxfk nxsvgxlysg crltxlrxek xgaatgx dai cxhxxpxkx glxxexlywe



17941 lsxltxxixe lgpytldrxx lyvngfhprx svpttstpgt stvhlatsgt psslpghtap  
 18001 vpllipftln ftitnlhyee nmqhpgerkf nttervlggl lgpmmfntsv gllysgcrlt  
 18061 llrpekngaa tgmdaicshr ldpkspglxx exlywelsxl txxixelgpy tldrxxslyvn  
 18121 gfthxxsxpt tstpgtstvx xgtsgtpssx pxxtxxxppl xpftxnxtit nlxxxxmxx  
 5 18181 pgsrkfntte xvlqglxpx fknxsvgxly sgcrltxlrx ekxgaatgxd aicxhxxpk  
 18241 xpglxxexly welsxltxxi xelgpytldr xslyvngfth qnsvpttstp gtstvywatt  
 18301 gtpssfpght epgpllipft fnftitnlhy eenmqhpgsr kfnttervlg glltpfknt  
 18361 svgplysgcr ltllrpekqe aatgvdtict hrvdpgpgl xxexlywels xltxxixelg  
 18421 pytldrxxsly vngfthxxsx ptstpgtst vxxgtsgtps sxpxttxxp lxpftxnx  
 10 18481 itnlxxxxm xpgsrkfnt texvlqglx pxfknxsvgx lysgcrltxl rxekxgaatg  
 18541 xdaicxhxx pxxpglxxex lywelsxltx xixelgpytl drxsllyvngf thrssvptts  
 18601 spgtstvhla tsqtpsslp htapvpplip ftlnftitnl hyeenmqhpg srkfntterv  
 18661 lqglkplfk stsvgplysg crltllrpek hgaatgvda clrlldptgp glxxexlywe  
 18721 lsxltxxixe lgpytldrxx lyvngfthxx sxpttstpgt stvxxgtsv psspxxtxx  
 15 18781 xpllxpftxn xtitnlxxx xmxpgsrkf nttervlggl lxpfxknxsv gxlysgcrlt  
 18841 xlrkekxgaa tgxdaicxh xpxkxpglxx exlywelsxl txxixelgpy tldrxxslyvn  
 18901 gfthrtsvpt tstpgtstvh latsgtpssl pghtapvppl ipftlnftit nlqeedmhr  
 18961 pgsrkfntte rvlqglspi fknxsvgply sgcrltslrp ekdgaatgmd avclyhpnk  
 19021 rpgldreqly celsqlthni telgpyaldr dslyvngfth qnsvpttstp gtstvywatt  
 20 19081 gtpssfpght xxxpllxpft xxtitnlxx xxxmxpgsr kfnttexvlg glxxpfxknx  
 19141 svxlysgcr ltxlrkekxg aatgxdaic hxxpxkxpgl xxexlywels xltxxixelg  
 19201 pytldrxxsly vngfthwssg lttstpwtst vdlgtsgtps pvpspttagp llvpftlnft  
 19261 itnlqyeedm hrpgsrkfna tervlglls pifkntsvgp lysgcrltll rpekqaatg  
 19321 vdticthrvd pigpglxxex lywelsxltx xixelgpytl drxsllyvngf thxxsxptts  
 25 19381 tpgtstvxg tsqtpssxp xtxxxpllxp ftxnxtitnl xxxmxpxg srkfnttexv  
 19441 lqglxpxfk nxsvgxlysg crltxlrxek xgaatgxdai cxhxxpxkx glxxexlywe  
 19501 lsxltxxixe lgpytldrxx lyvngfthrs fglttstpwt stvdlgtsgt pepvpstta  
 19561 gpplvpftln ftitnlqye dmhrpgsrkf nttervlggl ltpfmxsv tlysgcrlt  
 19621 llrpekdgaa trvdavcthr pdpkspglxx exlywelsxl txxixelgpy tldrxxslyvn  
 30 19681 gfthxxsxpt tstpgtstvx xgtsgtpssx pxxtxxxppl xpftxnxtit nlxxxxmxx  
 19741 pgsrkfntte xvlqglxpx fknxsvgxly sgcrltxlrx ekxgaatgxd aicxhxxpk  
 19801 xpglxxexly welsxltxxi xelgpytldr xslyvngfth wipvptstsp gtstvdlgsg  
 19861 tpsslpeptt agpllvpftl nftitnlqy edmghpgsrkf fnttervlgg llgpifknts  
 19921 vgplysgcr ltlrsekda atgvdaic hldpkspglx xexlywelsx ltxixelgp  
 35 19981 ytlrxxslyv ngfthxxsxp ttstpgtstv xxtsgtpss pxxtxxxppl lxpftxnx  
 20041 tnlxxxxmxx xpgsrkfntt exvlqglxp xfkxsvgxly ysgcrltxlr xekxgaatg  
 20101 daicxhxxp kxpglxxex lywelsxltx xixelgpytldr xslyvngfth hqfapntst  
 20161 pgtstvdltg sgtpslpsp tsagpllvpf tlnftitnlq yeedmhpgs rkfntterv  
 20221 qglgpmfkn tsvgllysgc rltllrpek gaatrvdavc thrpdpkspg lxxexlywel  
 40 20281 sxltxxixel gpytldrxxs lyvngfthxxs xpttstpgts vxxgtsagtp sxxpxxtapv  
 20341 pllipftlnf titnlhyeen mqhpgsrkfnt ttervlggl kplfkstsvg plysgcrltl  
 20401 lrpekhaat gvdaictlrl dptgpgldre rlywelsqlt nevtelegpyt ldrdalyvng  
 20461 ftqrssvptt sipgtsavhl etsgtpaslp ghtapgpplv pftlnftitn lqyeedmhrp  
 20521 gsrkfntter vlgglkplf kstsvgplys gcrltllrpe krgaatgvdt ictlrdpln  
 45 20581 pglrdreqlyw elskltrgi elgpylldrg slyvngfthr nfvpitstpg tstvhlgtse  
 20641 tpsslprpiv pgpllvpftl nftitnlqye eamrhpgrkf fnttervlgg llrplfknts  
 20701 igplysscr ltlrpekda atrvdaicth hpdpgspgln reqlywelsq lthgitelgp  
 20761 ytlrdslyv dgfthwspip ttstpgtsiv nlgtsgipps lpettxxopl lxpftxnx  
 20821 tnlxxxxmxx xpgsrkfntt ervlgglkp lfkstsvgpl ysgcrltllr pekdgvatrv  
 50 20881 daicthrpdp kipgldrqql ywelsqlths itelgpytldr rdslyvngfth qrsvpttst  
 20941 pgtftvqpet setpslpgp tatgpvllpf tlnftitnlq yeedmhpgs rkfntterv  
 21001 qglmplfk tsvslysgc rltllrpek gaatrvdavc thrpdpkspg ldrerlywk  
 21061 sqllthgitel gpytldrxxs lyvngfthqss mtttrtpdts tmhlatsrtp aslsgptas  
 55 21121 pllvftinf titnlryeen mhhpgsrkfnt ttervlggl rpvfkntsvg plysgcrltl  
 21181 lrpkkdgaat kvdaictyrp dpkspgldre qlywelsqlt hsitelgpyt ldrdalyvng  
 21241 ftqrssvptt sipgtptvdl gtsqtpvskp gpsaaspplv lftlnftitn lryeenmqh  
 21301 gsrkfntter vlgglrslf kstsvgplys gcrltllrpe kdgtatgvda ictlrdpks  
 21361 prldreqlyw elskltrnit elghyaldnd slfvngfthr ssvsttstpg tptvylgask  
 21421 tpasifgpa ashllilftl nftitnlrye enmwpgsrkf nttervlggl lrpplfknts  
 60 21481 gplysgsr ltlrpekdgaa tgvdacithr pdptgpgldr eqlylelsq lthsitelgpy  
 21541 tldrdslyvn gfthrsvpt tstgvvseep ftlnftinnl rymadmqqg slkfntitdnv

21601 mkhlslplfq rslgarytg crvialrsvk ngaetrvdll ctylqplsgp glpikqvfhc  
 21661 lsqqtghitr lgpysldkds lylnygnepg ldeppthpkp attflpplse attamgyhik  
 21721 tltlnftian lqyspdmkgk satfnstegv lqhlrlplfq kssmgyfylg cqlislrpek  
 21781 dgaatgvdtc ctvhpdpvpg gldiqqlwe lsqllthgvtg lgfyvldrds lfingyapqn  
 5 21841 lsirgeyqin fhivnwnlsn pdptsseyit llrdiqdkvt tlykgsqldh tfrfclvtnl  
 21901 tmdsvlvtvk alfssnldps lveqvfdkt lnasfhwlg tyqlvdihvt emessvyqpt  
 21961 sssstqhfyl nftitnlpye qdkagpggtn yqrnkknied alnqlfrnss ikayfsdcqv  
 22021 stfrsvpnrh htgvdslcnf splarrvdrv aiyeeflrmr rngtqlqntf ldrssvldvg  
 22081 yspnrnepln gnsdlpfwav iliglagllg litcliegvl vttrrrkkee eynvqqcpg  
 10 22141 yyqshldled lq

## SEQ ID NO. 2

15 CA125 nucleic acid Genbank No. AF414442

CDS 205..66663

1 aagcgttgca caattccccc aacctccata catacggcag ctcttctaga cacaggtttt  
 61 ccaggtcaa atgcggggac ccagccata tctcccacc tgagaaattt tggagtttca  
 121 gggagctcag aagctctgca gaggccaccc tctctgagg gattcttctt agacctccat  
 20 181 ccagaggcaa atgttgacct gtccatgctg aaaccctcag gccttctctg gtcattctct  
 241 cccacccgct ccttgatgac agggagcagg agcactaaag ccacaccaga aatggattca  
 301 ggactgacag gagccacctt gtcacctaa acatctacag gtgcaatcgt ggtgacagaa  
 361 catactctgc cctttacttc ccagataaag accttgcca gtcctacatc ttgggttg  
 421 ggaagaacca cccagtcttt ggggtgatg tctctgctc tccctgagtc aacctctaga  
 25 481 ggaatgacac actccgagca aagaaccagc ccatcgctga gtcccagggt caatggaact  
 541 cctcttagga actacctgac tacaagcatg gtttcaggat tgagtcccc aaggaccagg  
 601 accagttcca cagaaggaaa ttttaccaaa gaagcatcta catacacact cactgtagag  
 661 accacaagtg gccagtcac tgagaagtac acagtcccca ctgagacctc aacaactgaa  
 721 ggtgacagca cagagacccc ctgggacaca agatatattc ctgtaaaaat cacatctcca  
 30 781 atgaaaacat ttgcagattc aactgcatcc aaggaaaatg cccagtgctc tatgactcca  
 841 gctgagacca cagttactga ctacataact ccaggaagga caaacccatc atttgggaca  
 901 ctttattctt ccttccttga cctatcacct aaagggaccc caaattccag aggtgaaaca  
 961 agcctggaac tgattctatc aaccactgga tatcccttct cctctcctga acctggctct  
 1021 gcaggacaca gcagaataag taccagtgcg cctttgtcat catctgcttc agttctcgat  
 35 1081 aataaaatat cagagaccag catattctca ggccagagtc tcacctcccc tctgtctcct  
 1141 ggggtgcccg aggccagagc cagcacaatg ccaactcag ctatcccttt ttccatgaca  
 1201 ctaagcaatg cagaaacaag tgccgaaagg gtcagaagca caatttctc tctggggact  
 1261 ccataaatat ccacaaagca gacagcagag actatcctta ccttccatgc cttcgctgag  
 1321 accatggata taccagcac ccacatagcc aagactttgg cttcagaatg gttgggaagt  
 40 1381 ccaggtaccc ttggtggcac cagcacttca gcgctgacaa ccacatctcc atctaccact  
 1441 ttagtctcag aggagacca caccatcac tccacgagt gaaaggaaac agaaggaaact  
 1501 ttgaatacat ctatgactcc acttgagacc tctgctcctg gagaagagtc cgaatgact  
 1561 gccaccttgg tccccactct aggttttaca actcttgaca gcaagatcag aagtcactct  
 1621 caggtctctt catcccacc aacaagagag ctcagaacca caggcagcac ctctgggagg  
 45 1681 cagagttcca gcacagtgc ccacgggagc tctgacatcc tgaggggcaac cacttccagc  
 1741 acctcaaaag catcatcatg gaccagtga agcacagctc agcaatttag tgaacccag  
 1801 cacacacagt ggggtggagc aagtcctagc atgaaaacag agagaccccc agcatcaacc  
 1861 agtgtggcag cccctatcac cactctgtt cctcagtgg tctctggctt caccacctg  
 1921 aagaccagct ccacaaaagg gatttggctt gaagaaacat ctgcagacac actcatcgga  
 50 1981 gaatccacag ctggcccaac caccatcag tttgctgttc ccactgggat ttcaatgaca  
 2041 ggaggcagca gcaccagggg aagccagggc acaacccacc tactaccag agccacagca  
 2101 tcacttgaga catccgcaga tttgactctg gccacgaac gtgtcccagt ctccgtgtct  
 2161 ccagcagtg gcaagacggc tggctgctca agtctccag gagggaacaa gtcactat  
 2221 acaatgggtt cttctgtcat ccctgagaca tcatctctac agtctcagc ttccagggaa  
 55 2281 ggaaccagcc tgggactgac tccattaaac actagacatc ccttctcttc cctgaacca  
 2341 gactctgcag gacacaccaa gataagcacc agcattcctc tgttgcate tgcctcagtt  
 2401 cttgaggata aagtgtcagc gaccagcaca ttctcacacc acaagccca ctcactatt  
 2461 accacaggga ctctgaaat ctcaacaaag acaagccca gctcagccgt tcttctctcc  
 2521 atgaccttaa gcaatgcagc aacaagtcct gaaagagtca gaaatgcaac tccccctctg

2581 actcatccat ctccatcagg ggaagagaca gcagggagtg tctctactct cagcacctct  
2641 gctgagacta cagactcacc taacatccac ccaactggga cactgacttc agaatcgtca  
2701 gagagtccta gcactctcag cctcccaagt gtctctggag tcaaaaccac attttcttca  
2761 tctactcctt ccactcatct atttactagt ggagaagaaa cagaggaaac ttcgaatcca  
5 2821 tctgtgtctc aacctgagac ttctgtttcc agagtaagga ccaccttggc cagcacctct  
2881 gtccctaccc cagtattccc caccatggac acctggccta caggttcagc tcagttctct  
2941 tcatcccacc tagtgagtga gctcagagct acgagcagta cctcagttac aaactcaact  
3001 ggttcagctc ttctctaaaat atctcacctc actgggacgg caacaatgtc acagaccaat  
3061 agagacacgt ttaatgactc tgcctgaccc caaagcacaa cttggccaga gactagtcct  
10 3121 agattcaaga cagggttacc ttccagcaaca accactgttt caacctctgc cacttctctc  
3181 tctgtactcg taatggtctc taaattcact totccagcaa ctagtcccat ggaagcaact  
3241 tctatcaggg aacctcaac aacctcctc acaacagaga ccacgaatgg ccaggtctct  
3301 atggctgtgg cttctaccaa catcccattt ggaaagggct acattactga aggaagattg  
3361 gacacaagcc atctgcccac tggaaaccaca gcttctctct agacatctat ggattttacc  
15 3421 atggccaaag aaagtgtctc aatgtcagta totccatctc agtccatgga tgcgtctggc  
3481 tcaagcactc caggaaggac aagccaattc gttgacacat tttctgatga tgtctatcat  
3541 ttaacatcca gagaattac aatacctaga gatggaacaa gctcagctct gactccacaa  
3601 atgactgcaa ctcacctctc atctcctgat cctggctctg ctagaagcac ctggcttggc  
3661 atcttgtcct catctccttc ttctcctact ccaaagtca caatgagctc cacattttca  
20 3721 actcagagag tcaccacaag catgataatg gacacagttg aaactagtct gtggaacatg  
3781 ccaacttac cttccacgac ttccctgaca ccaagtaata ttccaacaag tgggtgccata  
3841 gaaaaaagca ccctgggttc cttggacact ccactctccag ccacatcatt ggaggcatca  
3901 gaagggggac ttccaacctc cagcacctac cctgaatcaa caaacacacc cagcatccac  
3961 ctggagcac acgctagtct agaaagtcca agcaccatca aacttaccat ggcttcagta  
25 4021 gtaaaacctg gctcttacac acctctcacc ttccctcaa tagagacca cattcatgta  
4081 tcaacagcca gaatggctta ctcttctggg tcttcacctg agatgacagc tcctggagag  
4141 actaacactg gtagtacctg ggacccccac acctacatca ccactacgga tcctaaggat  
4201 acaagttcag ctcaggtctc tacaccccac tcagttagga cactcagaac cacagaaaac  
4261 catccaaaga cagagtccgc caccctcagc gcttactctg gaagtcctaa aatctcaagt  
30 4321 tcaccaatc tcaccagtc ccgccacaaa gcattggacca tcacagacac aactgaacac  
4381 tccactcaat tacattacac aaaattggca gaaaaatcat ctggatttga gacacagtca  
4441 gctccaggac ctgtctctgt agtaatccct acctccccta ccattggaag cagcacattg  
4501 gaactaactt ctgatgtccc aggggaaccc ctggctcctt ctcccagtgga gcagaccaca  
4561 atcactctcc ccattggcaac atggctgagt accagtttga cagaggaaat ggcttcaaca  
35 4621 gaccttgata tttcaagtcc aagttcaccc atgagtacat ttgctatttt tccacctatg  
4681 tccacacett ctcattgaact ttcaaagtca gaggcagata ccagtgccat tagaaatata  
4741 gattcaacaa cgttggatca gcacctagga atcaggagtt tgggcagaaac tggggactta  
4801 acaactgttc ctatcacccc actgacaacc acgtggacca gttgtattga acactcaaca  
4861 caagcacagg acaccttttc tgcacagatg agtccctact acgtgacaca gtactcaaaa  
40 4921 gatcaaacat ctataccagc ctcagcatcc ccttcccato ttactgaagt ctacctgag  
4981 ctggggacac aagggagaag ctctctgag gcaaccaact ttggaaac atctacagac  
5041 acactgtcca gagagattga gactggccca acaaacatto aatccactcc acccatggac  
5101 aacacaacaa cagggagcag tagtagtgga gtcaccctgg gcattagccca ccttcccata  
5161 ggaacatcct cccagctga gacatccaca aacatggcac tggaaagaag aagttotaca  
45 5221 gccactgtct ctatggctgg gacaatggga ctcttgtta ctagtgtctc aggaagaagc  
5281 atcagccagt cattaggaag agtttctctc gtcttctctg agtcaactac tgaaggagt  
5341 acagattcta gtaagggaag cagcccaagg ctgaacacac agggaaatac agctctctcc  
5401 tctctctctg aacctcagta tgcctgaagga agccagatga gcacaagcat ccctctaacc  
5461 tcatctccta caactcctga tgtggaattc atagggggca gcacattttg gaccaaggag  
50 5521 gtcaccacag ttatgaacct agacatctcc aagtcttcag caaggacaga gtcacgctca  
5581 gctaccttta tgtccacagc tttgggaagc actgaaaata caggaaaaga aaaactcaga  
5641 actgctctca tggatcttcc atctccaact ccactcaatgg aggtgacacc atggatttct  
5701 ctcactctca gtaatgcccc caataccaca gattcacttg acctcagcca tggggtgcac  
5761 accagctctg cagggacttt ggccactgac aggtcattga atactggtgt cactagagcc  
55 5821 tccagattgg aaaacggctc tgatacctct tctaagctcc tgtctatggg aaacagact  
5881 cacacttcca tgactgacac agagaagagt gaagtgtctt ctccaatcca tccccgacct  
5941 gagacctcag ctcttgagc agagaccact ttgacttcca ctcttgaaa cagggccata  
6001 agcttaacat tgccttttcc atccattcca gtggaagaag tcaattotac aggcataacc  
6061 tcaggaccag acatcaactc agcacccatg acacattctc ccatcaccac accaacaatt  
60 6121 gtatggacca gtacaggcac aattgaacag tccactcaac cactacatgc agtttcttca  
6181 gaaaaagttt ctgtgcagac acagtcaact ccatatgtca actctgtggc agtgcctgct

5 6241 tccccctacc atgagaattc agtctctctt ggaagcagca catcctctcc atattcctca  
 6301 gcctcacttg aatccttgga ttccacaatc agtaggagga atgcaatcac ttccctggcta  
 6361 tgggacctca ctacatctct cccactaca acttggtcaa gtactagtgt atctgaggca  
 6421 ctgtcctcag gccattctgg ggtttcaaac ccaagttaa ctacgactga atttccactc  
 6481 ttttcagctg catccacatc tgcctgtaag caaagaaatc cagaacaga gacctatggt  
 6541 cccagaata cagccgcag tactttgaac actgatgcat cctcgggtcac aggtctttct  
 6601 gagactcctg tgggggcaag tatcagctct gaagtccctc ttccaatggc cataacttct  
 6661 agatcagatg tttctggcct tacatctgag agtactgcta acccgagttt aggcacagcc  
 6721 tcttcagcag ggaccaaatt aactaggaca atatccctgc ccacttcaga gtctttgggt  
 10 6781 tccttttaga tgaacaagga tccatggaca gtgtcaatcc ctttgggggt ccacccaact  
 6841 actaatacag aaacaagcat cccagtaaag agcgaggtgc cacttggtt gtccacagta  
 6901 gcatcagatg taattgacac accttcagat ggggctgaga gtattccac tgtctccttt  
 6961 tccccctccc ctgatactga agtgacaact atctcacatt tcccagaaaa gacaactcat  
 7021 tcattttaga ccatttcac tctcactcat gaggttgact caagagtgc acctattcct  
 15 7081 ggggattgga tgagttcagc tatgtctaca aagcccacag gagccagtc ctcattaca  
 7141 ctgggagaga gaaggacaat cactctgct gctccaacca cttccctcat agttctcact  
 7201 gctagtttca cagagaccag cacagtttca ctggataatg aaactacagt aaaaacctca  
 7261 gatattcctg acgcacggaa acaaatgag ctccccctc atagcagttc ttctctctgat  
 7321 ctgatcaaca cctccatagc ttcttcaact atggatgtca ctaaacacgc tccatcagt  
 20 7381 cccactagca tctcaggaat gacagcaagt tctctcccat ctctcttctc tccagataga  
 7441 cccaggttc ccacatctac aacagagaca aatacagcca cctctccatc tgtttccagt  
 7501 aacacctatt ctcttgatgg gggctccaat gtgggtggga ctccatccac ttaccaccc  
 7561 tttacaatca cccacctgt cgagacaagc tcggccctat tagcctggtc tagaccagta  
 25 7621 agaactttca gcacctggt cagcactgac actgcctccg gagaataatc taccttagc  
 7681 aattctgtgg tgacttctgt tccagcacca ggtacatggg ccagtgtagg cagtactact  
 7741 gacttacctg ccattgggct tctcaagaca agtccctgc gagaggcaca ctacttcta  
 7801 gcatcaacta ttgaaccagc cactgccttc actcccatc tctcagcagc agtggtcact  
 7861 ggatccagtg ctacatcaga agccagtctt ctactacga gtgaaagcaa agccattcat  
 7921 tcttcaccac agaccccaac tacaccaccc tctggagcaa actgggaaac ttcagctact  
 30 7981 cctgagagcc ttttggtagt cactgagact tcagacacaa cacttacctc aaagattttg  
 8041 gtcacagata ccattctgtt ttcaactgtg tccacgcac cttctaaatt tccaagtacg  
 8101 gggactctgt ctggagcttc ctccctact ttactcccgg aactccagc catccctctc  
 8161 actgccactg agccaacaag ttacttagct acatcctttg attccaccac actggtgact  
 35 8221 atagcttcgg atagctctgg cacagtcaca gagactaccc tgaccatgtc agagacctca  
 8281 aatggtgatg cactggttct taagacagta agtaaccac ataggagcat ccctggaatc  
 8341 actatccaag gagtaacaga aagtccactc catccttctt ccacttcccc ctctaagatt  
 8401 gttgctccac ggaatacaac ctatgaaggt tcgatcacag tggcactttc tactttgctt  
 8461 gcgggaacta ctgggtccct tgtattcagt cagagttctg aaaactcaga gacaacggct  
 40 8521 ttggtagact catcagctgg gcttgagagg gcatctgtga tgccactaac cacaggaagc  
 8581 cagggtatgg ctagctctgg aggaatcaga agtgggtcca ctactcaac tggaaacaaa  
 8641 acattttctt ctctccctct caccatgaac ccaggtgagg ttacagccat gtctgaaatc  
 8701 accacgaaca gactgacagc tactcaatca acagcaccca aagggatacc tgtgaagccc  
 8761 accagtgtct agtcaggcct cctaaccct gtctctgcct cctcaagccc atcaaaggcc  
 8821 tttgcctcac tgactacagc tccccatca acttggggga tcccacagtc taccttgaca  
 45 8881 tttgagtttt ctgaggtccc aagtttggat actaagtccg ctctcttacc aactcctgga  
 8941 cagtccctga acaccattcc agactcagat gcaagcacag catcttcccc actgtccaag  
 9001 tctccagaaa aaaacccaag ggcaaggatg atgacttcca caaaggccat aagtgcaagc  
 9061 tcattttcaat caacaggttt tactgaaacc cctgagggat ctgcctcccc ttctatggca  
 50 9121 gggcatgaac ccagagtcct cacttcagga acaggggacc ctgatatgac ctcagagagc  
 9181 atgtcttatc cagacccaag caaggcatca tcagctatga catcgacctc tcttgcatca  
 9241 aaactcacia ctctcttcag cacaggtcaa gcagcaaggt ctgggtctag ttctctctcc  
 9301 ataagcctat ccactgagaa agaaacaagc ttcttttccc ccactgcac cactccaga  
 9361 aagacttcac tatttcttgg gccttccatg gcaaggcagc ccaacatatt ggtgcatctt  
 9421 cagacttcag ctctgacact ttctccaaca tccactctaa atatgtccca ggaggagcct  
 55 9481 cctgagttaa cctcaagcca gaccattgca gaagaagagg gaacaacagc tgaaacacag  
 9541 acgttaacct tcacaccatc tgagacccca acatccttgt tacctgtctc ttctccaca  
 9601 gaaccacag ccagaagaaa gaggttctcca gaaacatggg caagctctat ttcatgtcct  
 9661 gccagacct ccttgggtga aacaactgat ggaacgctag tgaccacct aaagatgtca  
 9721 agccaggcag cacaaggaaa ttccacgtgg cctgccccag cagaggagac ggggaccagt  
 60 9781 ccagcaggca catccccagg aagcccagaa gtgtctacca ctctcaaat catgagctcc  
 9841 aaggaaccca gcatcagccc agagatcagg tccactgtgc gaaattctcc ttggaagact

9901 ccagaaacaa ctgttcccat ggagaccaca gtggaaccag tcacccttca gtccacagcc  
 9961 ctaggaagtg gcagcaccag catctctcac ctgcccacag gaaccacatc accaaccaag  
 10021 tcaccaacag aaaatatgtt ggctacagaa aggggtctccc tctccccatc cccacctgag  
 10081 gcttggacca acctttattc tggaaactcca ggagggacca ggcagtcact ggccacaatg  
 5 10141 tcctctgtct ccctagagtc accaactgct agaagcatca caggggactgg tcagcaaaagc  
 10201 agtccagaac tgggtttcaaa gacaactgga atggaattct ctatgtggca tggctctact  
 10261 ggagggacca caggggacac acatgtctct ctgagcacat cttccaatat ccttgaagac  
 10321 cctgtaacca gccc aaactc tgtgagctca ttgacagata aatccaaaca taaaaccgag  
 10381 acatgggtaa gcaccacagc catctccctcc actgtcctga ataataagat aatggcagct  
 10 10441 gaacaacaga caagtgcgac tgtggatgag gcttattcat caactagtcc ttggtcagat  
 10501 cagacatctg ggagtgcacat cacccttggg gcactcctcg atgtcacaaa cacattatac  
 10561 atcacctcca cagcacaac cactcacta gtgtctctgc cctctggaga ccaaggcatt  
 10621 acaagcctca ccaatccctc agggagaaaa acaagctctg cgtcatctgt cacatctcct  
 10681 tcaatagggc ttgagactct gaggggcaat gtaagtgcag tgaaaagtga cattgcccct  
 15 10741 actgctgggc atctatctca gacttcatct cctgcggaag tgagcatctc ggacgtaaac  
 10801 acagctccta ctccaggat ctccaccacc atcaccacca tgggaaccaa ctcaatctca  
 10861 actaccacac ccaaccacaga agtgggtatg agtaccatgg acagcaccac ggccacagag  
 10921 aggcgcacaa cttctacaga acacccttcc acctgggtctt ccacagctgc atcagattcc  
 10981 tggactgtca cagacatgac ttcaaaccttg aaagtgcgaa gatctcctgg aacaatttcc  
 20 11041 acaatgcata caacttcatt cttagcctca agcactgaat tagactccat gtctactccc  
 11101 catggcgtga taactgtcat tggaaaccagc ctggtcactc catcctctga tgcctcagct  
 11161 gtaaaagacag agaccagtac aagtgaaga acattgagtc ctccagacac aactgcattct  
 11221 actcccatct caactttttc tctgttccag aggatgagca tctcagttcc tgacatttta  
 25 11281 agtacaagtt ggactcccag tagtacagaa gcagaagatg tgccctgttc aatggtttct  
 11341 acagatcatg ctagtacaaa gactgaccca aatacgcccc tgtccacttt tctgtttgat  
 11401 tctctgtcca ctcttgactg ggacactggg agatctctgt catcagccac agccactacc  
 11461 tcagctcctc agggggccac aactccccag gaactcactt tggaaaccat gatcagccca  
 11521 gctacctcac agttgccctt ctctataggg cacattacaa gtgcagtcac accagctgca  
 11581 atggcaagga gctctggagt tactttttca agaccagatc ccacaagcaa aaaggcagag  
 30 11641 cagacttcca ctacagctcc caccaccact tctgcacatc cagggcaggt gccagatca  
 11701 gcagcaacaa ctctggatgt gatccacac acagcaaaaa ctccagatgc aacttttcag  
 11761 agacaagggc agacagctct tacaacagag gcaagagcta catctgactc ttggaatgag  
 11821 aaagaaaaat caaccccaag tgcaccttgg atcactgaga tgatgaattc tgtctcagaa  
 11881 gataccatca aggaggttac cagctcctcc agtgtattaa aggaccctga atacgctgga  
 35 11941 cataaacttg gaatctggga cgacttcac ccaagtttg gaaaagcagc ccatatgaga  
 12001 gagttgcccc ttctgagtc accacaggac aaagaggcaa ttcacccttc tacaacaca  
 12061 gtagagacca caggctgggt cacaagttcc gaacatgctt ctcatccac tatccagcc  
 12121 cactcagcgt catccaaact cacatctcca gtggttaca cctccaccag ggaacaagca  
 12181 atagtttcta tgtcaacaac cacatggcca gactctacaa gggctagaac agagccta  
 40 12241 tccttcttga ctattgaact gaggagctc agcccttaca tggacaccag ctcaaccaca  
 12301 caaacaagta ttatctcttc cccaggttcc actgcgatca ccaaggggcc tagaacagaa  
 12361 attacctcct ctaagagaat atccagctca ttcttggccc agtctatgag gtcgtcagac  
 12421 agccctcag aagccatcac caggctgtct aactttctg ccatgacaga atctggagga  
 12481 atgaccttg ctatgcaaac aagtccacct ggcgctacat cactaagtgc acctactttg  
 45 12541 gatacatcag ccacagcctc ctggacaggg actccactgg ctacgactca gagatttaca  
 12601 tactcagaga agaccactct ctttagcaaa ggtcctgagg atacatcaca gccaaagcct  
 12661 ccctctgttg aagaaaccag ctcttctct tccctggtac ctatccatgc tacaacctcg  
 12721 ccttccaata ttttgttgac atcacaaggg cacagtcctt cctctactcc acctgtgacc  
 12781 tcagttttct tgtctgagac ctctggcctg gggaagacca cagacatgtc gaggataagc  
 50 12841 ttggaacctg gcacaagttt acctcccaat ttgagcagta cagcaggtga ggcgttatcc  
 12901 acttatgaag cctccagaga tacaaggca attcatcatt ctgcagacac agcagtgacg  
 12961 aatatggagg caaccagtcc tgaatattct cctatcccag gccatacaaa gccatccaaa  
 13021 gccacatctc cattgggttac ctccacatc atgggggaca tcaacttctc cacatcagta  
 13081 tttggctcct ccgagaccac agagattgag acagtgtcct ctgtgaacca gggacttcag  
 55 13141 gagagaagca catcccaggt ggcagctct gctacagaga caagcactgt cattaccat  
 13201 gtgtctagtg gtgatgtac tactcatgtc accaagacac aagccacttt ctctagcgga  
 13261 acatccatct caagccctca tcagtttata acttctacca acacatttac agatgtgagc  
 13321 accaaccctt ccacctctct gataatgaca gaatcttcag gagtgacct caccacccaa  
 13381 acaggtccta ctggagctgc aacacagggt ccatatctct tggacacatc aacctgcct  
 60 13441 tacttgacag agactccatt agctgtgact ccagatttta tgcaatcaga gaagaccact  
 13501 ctcataagca aaggtcccaa ggtgtgacc tggacaagcc ctccctctgt gcgagaacc

	13561	agctatccct	cttccctgac	acctttcttg	gtcacaacca	tacctcctgc	cacttccacg
	13621	ttacaagggc	aacatacatc	ctctcctgtt	tctgcgactt	cagttcttac	ctctggactg
	13681	gtgaagacca	cagatatgtt	gaacacaagc	atggaacctg	tgaccaattc	acctcaaaat
	13741	ttgaacaatc	catcaaatga	gatactggcc	acttttggcag	ccaccacaga	tatagagact
5	13801	attcatcctt	ccataaacia	agcagtgaac	aatatgggga	ctgccagttc	agcacatgta
	13861	ctgcattcca	ctctcccagt	cagctcagaa	ccatctacag	ccacatctcc	aatggttctt
	13921	gcctccagca	tgggggacgc	tcttgcttct	atatcaatac	ctggttctga	gaccacagac
	13981	attgagggag	agccaacatc	ctccctgact	gctggacgaa	aagagaacag	caccctccag
	14041	gagatgaact	caactacaga	gtcaaacatc	atcctctcca	atgtgtctgt	gggggctatt
10	14101	actgaagcca	caaaaatgga	agtccccctt	tttgatgcaa	cattcatacc	aactcctgct
	14161	cagtoacaaa	agttcccaga	tattttctca	gtagccagca	gtagacttcc	aaactctcct
	14221	cccatgacaa	tatctaccca	catgaccacc	accagacag	ggctctctgg	agctacatca
	14281	aagattccac	ttgccttaga	cacatcaacc	ttggaaacct	cagcagggac	tccatcagtg
	14341	gtgactgagg	ggtttgccca	ctcaaaaata	accactgcaa	tgaacaatga	tgtcaaggac
15	14401	gtgtcacaga	caaaccctcc	ctttcaggat	gaagccagct	ctccctcttc	tcaagcacct
	14461	gtccttgta	caaccttacc	ttcttctgtt	gctttcacac	cgcaatggca	cagtacctcc
	14521	tctcctgttt	ctatgtcctc	agttcttact	tcttcaactg	taaagaccgc	aggcaagggtg
	14581	gatacaagct	tagaaacagt	gaccagttca	cctcaaagta	tgagcaaac	tttggaatgac
	14641	atatcggtca	cttcagcagc	caccacagat	atagagacaa	cgcattcctc	cataaacaca
20	14701	gtagttaaca	atgtggggac	cacoggttca	gcatttgaat	cacattctac	tgtctcagct
	14761	taccagagc	catctaaagt	cacatctcca	aatgttacca	cctccaccat	ggaagacacc
	14821	acaatttccc	gatcaatacc	taaatcctct	aagactacaa	gaactgagac	tgagacaact
	14881	tcctccctga	ctcctaaact	gaggagagac	agcatctccc	aggagatcac	ctcgtccaca
	14941	gagacaagca	ctgttcttta	caaagagctc	actggtgcca	ctaccgaggt	ttccaggaca
25	15001	gatgtcactt	cctctagcag	tacatccttc	cctggccctg	atcagtcac	agtgtacta
	15061	gacatctcca	cagaaaccaa	caccaggctg	tctacctccc	caataatgac	agaatctgca
	15121	gaaataacca	tcaccaccca	aacaggctct	catggggcta	catcacagga	tacttttacc
	15181	atggaccat	caaatacaac	ccccaggcca	gggatccact	cagctatgac	tcatggattt
	15241	tcacaattgg	atgtgaccac	tcttatgagc	agaattccac	aggatgtatc	atggacaagt
30	15301	cctccctctg	tggataaaac	cagctccccc	tcttctcttc	tgtcctcacc	tgcaatgacc
	15361	acaccttccc	tgatttcttc	taccttacca	gaggataagc	tctcctctcc	tatgacttca
	15421	cttctcacct	ctggcctagt	gaagattaca	gacatattac	gtacacgctt	ggaacctgtg
	15481	accagctcac	ttccaaattt	cagcagcacc	tcagataaga	tactggccac	ttctaaagac
	15541	agtaaagaca	caaaggaaat	tttctcttct	ataaacacag	aagagaccaa	tgtgaaagcc
35	15601	acaactcaga	gacatgaatc	ccattccctc	gcactggctg	actcagagac	acccaagacc
	15661	acaactcaga	tgggttatcac	caccactgtg	ggagatccag	ctccttccac	atcaatgcca
	15721	gtgcatgggt	cctctgagac	tacaaacatt	aagagagagc	caacatattt	cttgactcct
	15781	agactgagag	agaccagtac	ctctcaggag	tccagctttc	ccacggacac	aagttttcta
	15841	ctttccaaag	tccccactgg	tactattact	gaggctctca	gtacaggggt	caactcttct
40	15901	agcaaaatth	ccaccccaga	ccatgataag	tccacagtgc	cacctgacac	cttcacagga
	15961	gagatcccca	gggtcttcac	ctcctctatt	aagacaaaat	ctgcagaaat	gacgatcacc
	16021	acccaagcaa	gtcctcctga	gtctgcatcg	cacagtaacc	ttcccttgga	cacatcaacc
	16081	acactttccc	agggagggac	tcattcaact	gtgactcagg	gattcccata	ctcagaggtg
	16141	accactctca	tgggcatggg	tcttgggaat	gtgtcatgga	tgacaactcc	ccctgtggaa
45	16201	gaaaccagct	ctgtgtcttc	cctgatgtct	tcacctgcca	tgacatcccc	ttctcctggt
	16261	tcctccacat	caccacagag	catccccctc	tctcctcttc	ctgtgactgc	acttctactt
	16321	tctgttctgg	tgacaaccac	agatgtgttg	ggcacaacaa	gcccagagtc	tgtaaccagt
	16381	tcacctccaa	atttgagcag	catcactcat	gagagaccgg	ccacttacaa	agacactgca
	16441	cacacagaag	ccgcatgca	tcattccaca	aacaccgcag	tgaccaatgt	agggacttcc
50	16501	gggtctggac	ataaatcaca	atcctctgtc	ctagctgact	cagagacatc	gaaagccaca
	16561	cctctgatga	gtaccacctc	caccctgggg	gacacaagtg	tttccacatc	aactccta
	16621	atctctcaga	ctaaccaaat	tcaaacagag	ccaacagcat	ccctgagccc	tagactgagg
	16681	gagagcagca	cgtctgagaa	gaccagctca	acaacagaga	caaatactgc	cttttcttat
	16741	gtgcccacag	gtgctattac	tcaggcctcc	agaacagaaa	tctcctctag	cagaacatcc
55	16801	atctcagacc	ttgatcgccc	cacaatagca	cccgacatct	ccacaggaaat	gatcaccagg
	16861	ctcttcacct	cccccatcat	gacaaaatct	gcagaaatga	ccgtcaccac	tcaaacact
	16921	actcctgggg	ctacatcaca	gggtatcctt	ccttggggaca	catcaaccac	acttttccag
	16981	ggagggactc	attcaaccgt	gtctcaggga	ttcccacact	cagagataac	cactctctgg
	17041	agcagaaccc	ctggagatgt	gtcatggatg	acaactcccc	ctgtgggaaga	aaccagctct
60	17101	gggttttccc	tgatgtcacc	ttccatgaca	tccccttctc	ctgtttctctc	cacatcacca
	17161	gagagcatcc	cctcctctcc	tctcctctgtg	actgcacttc	ttacttctgt	tctggtgaca

5 17221 accaccaatg tattggggac aacaagccca gagacogtaa cgagttcacc tccaaattta  
 17281 agcagcccca cacaggagag actgaccact tacaaagaca ctgcgcacac agaagccatg  
 17341 catgcttcca tgcatacaaa cactgcagtg gccaacgtcg ggacctccat ttctggacat  
 17401 gaatcacaaat cttctgtccc agctgattca cacacatcca aagccacatc tccaatgggt  
 17461 atcaccttcg ccattggggga tacaagtgtt tctacatcaa ctctgtcctt ctttgagact  
 17521 agaattcaga ctgaatcaac atcctctttg attccttgat taaggagacc caggagctct  
 17581 gaggagatca acactgtgac agagaccagc actgtccttt cagaagtgc cactactact  
 17641 actactgagg tctccaggac agaagtatc acttccagca gaacaacatc ctcagggcct  
 17701 gatcattcca aaatgtcacc ctacatctcc acagaaacca tcaccaggct ctccactttt  
 10 17761 ccttttgtaa caggatccac agaatggcc atcaccaacc aaacaggtcc tatagggact  
 17821 atctcacagg ctacccttac cctggacaca tcaagcacag ctctctggga agggactcac  
 17881 tcacctgtga ctacagagatt tccacactca gaggagacca ctactatgag cagaagtact  
 17941 aagggcggtg catggcaaaag cctccctct gtggaagaaa ccagttctcc ttcttcccca  
 18001 gtgcctttac ctgcaataac ctcacattca tctctttatt ccgcagatc aggaagttagc  
 15 18061 ccactttctg ctctcctgt gacttccctt ctcacctctg gcaggaggaa gaccatagac  
 18121 atgttgagaca cacactcaga actgttgacc agctccttac caagtgcagg tagcttctca  
 18181 ggtgagatac tcacttctga agcctccaca aatacagaga caattcactt ttcagagaac  
 18241 acagcagaaa ccaatatggg gaccaccaat tctatgcata aactacattc ctctgtctca  
 18301 atccactccc agccatccgg acacacacct ccaaagggtta ctggatctat gatggaggac  
 20 18361 gctattgttt ccacatcaac acctgggttct cctgagacta aaaatgttga cagagactca  
 18421 acatcccttc tgactcctga actgaaagag gacagcacog ccctgtgtat gaactcaact  
 18481 acagagtcaa acactgtttt ctccagtgtg tccctggatg ctgctactga ggtctccagg  
 18541 gcagaagtca cctactatga tctacattc atgcccagtt ctgctcagtc aacaaagtcc  
 18601 ccagacattt cacttgaagc cagcagcagt cattctaact ctctcctct gacaatatct  
 25 18661 acacacaaga ccatcgccac acaaacaggt ccttctgggg tgacatctct tggccaactg  
 18721 accctggaca catcaaccat agccacctca gcaggaactc catcagccag aactcaggat  
 18781 tttgttagatt cagaacaac cagtgtcatg aacaatgat tcaatgatgt gttgaagaca  
 18841 agccctttct ctgcagaaga agccaaactc ctctcttctc aggcacctct ccttgtgaca  
 30 18901 acctcacctt ctctgtatc tccacattg caagagcaca gtacctctc tcttgtttct  
 18961 gtgacctcag taccacccc tactctggcg aagatcacag acatggacac aaacttagaa  
 19021 cctgtgactc gttcacctca aaatttaagg aacaccttgg ccacttcaga agccaccaca  
 19081 gatacacaca caatgcatcc ttctataaac acagcaatgg ccaatgtggg gaccaccagt  
 19141 tcaccaaattg aattctatct tactgtctca cctgactcag acccatataa agcccaactc  
 35 19201 gcagtagtta tcacttccac ctctgggggac tcaatagttt ccacatcaat gcctagatcc  
 19261 tctgcgatga aaaagattga gtctgagaca actttctccc tgatathtag actgaggagg  
 19321 actagcacct ccagaaaaat tggctcatcc tcagacacaa gcacggtctt tgacaaagca  
 19381 ttactgtctg ctactactga ggtctccaga acagaaactc cctctctag cagaacatcc  
 19441 atccaaggca ctgaaaagcc cacaatgtca ccggacacct ccacaagatc tgtcaccatg  
 40 19501 ctttctactt ttgctggcct gacaaaatcc gaagaaagga ccattgccac ccaaacaggt  
 19561 cctcataggg cgacatcaca gggtaacctt acctgggaca catcaatcac aacctcacag  
 19621 gcagggaccc actcagctat gactcatgga ttttcacaa tagatttgc cactottagc  
 19681 agtagagttc ctgagtagat atcaggggaca agccacctc ctgtggaaaa aaccagctct  
 19741 tctcttctcc ttctgtcttt accagcaata acctcacctg cccctgtacc tactacatta  
 45 19801 ccagaaagta ggccgtcttc tctgttctat ctgacttccac tccccacctc tggcctagtg  
 19861 aagaccacag atatgtggc atctgtggcc agtttacctc caaacttggg cagcacctca  
 19921 cataagatac cgactacttc agaagacatt aaagatacag agaaaatgta tccttccaca  
 19981 aacatagcag taaccaatgt ggggaccacc acttctgaaa aggaatctta ttctgtctgc  
 20041 ccagcctact cagaaccacc caaagtcacc tctccaatgg ttacctctt caacataagg  
 50 20101 gacaccattg ttccacatc catgctggc tctctgaga ttacaaggat tgagatggag  
 20161 tcaacattct cgttggtcca tgggctgaag ggaaccagca cctcccagga ccccatcgta  
 20221 tccacagaga aaagtgtgt ttgacacaa ttgacctggt gtgctactga gaccttagg  
 20281 acagaagtgt cctcttctag aagaacatcc attccaggcc ctgatcttc cacagagtca  
 20341 ccagacatct ccactgaagt gatccccagc ctgcctatct cccttggcat tacagaatct  
 55 20401 tcaaatatga coatcatcac tcgaacaggt cctctctgtg gctctacatc acagggcaca  
 20461 tttaccttgg acacaccaac tacatctccc agggcaggaa cacactcag ggcactcag  
 20521 gaatttccac actcagaaat gacctgtgc atgaacaagg acctgagat tctatcatgg  
 20581 acaatccctc cttctataga gaaaaccagc ttctctctt ccctgatgcc ttcaccagcc  
 20641 atgacttcac ctctgtttc ctcaacatta ccaaagacca ttcacaccac tcttctctc  
 20701 atgacctcac tgctcacccc tagcctagt atgaccacag acacattggg cacaagccca  
 60 20761 gaacctacaa ccagttcacc tccaaatttg agcagtacct cacatgtgat actgacaaca  
 20821 gatgaagaca ccacagctat agaagccatg catccttcca caagcacagc agcgactaat

20881 gtggaaacca cctgttctgg acatgggtca caatcctctg tccctaactga ctcagaaaaa  
20941 accaaggcca cagctccaat ggataccacc tccaccatgg ggcatacaac tgtttccaca  
21001 tcaatgtctg tttcctctga gactacaaaa attaagagag agtcaacata ttccttgact  
21061 cctgggactga gagagaccag cattiecccaa aatgccagct tttccactga cacaagtatt  
5 21121 gttctttcag aagtccccac tgggtactact gctgaggtct ccaggacaga agtcacctcc  
21181 tctggtagaa catccatccc tggcccttct cagtccacag ttttgccaga aatatccaca  
21241 agaacaatga caaggctctt tgctcgcgcc accatgacag aatcagcaga aatgacctc  
21301 cccactcaaa caggctcctt tgggtctacc tcacaggata cccttacctt ggacacatcc  
21361 accacaaagt cccaggcaaa gactcattca actttgactc agagatttcc aactcagag  
10 21421 atgaccactc tcatgagcag aggtcctgga gatatgtcat ggcaaagctc tccctctctg  
21481 gaaaatccca gctctctccc ttccctgctg tctttacctg ccacaacctc acctcctccc  
21541 atttctccca cattaccagt gactatctcc tccctctctc ttctgtgtac ttcacttctc  
21601 acctctagcc cggtaacgac cacagacatg ttacacacaa gcccagaact tgtaaccagt  
21661 tcacctccaa agctgagcca cacttcagat gagagactga ccactggcaa ggacaccaca  
15 21721 aatacagaag ctgtgcattc ttccacaaac acagcagcgt ccaatgtgga gattcccagc  
21781 tttggacatg aatccccctc ctctgcctta gctgactcag agacatccaa agccacatca  
21841 ccaatgttta ttacctccac ccaggaggat acaactgttg ccatatcaac ccctcacttc  
21901 ttggagacta gcagaattca gaaagagtca atttctccc tgagccctaa attgaggag  
21961 acaggcagtt ctgtggagac aagctcagcc atagagacaa gtgtgtctct tctggaagt  
20 22021 tccattggtg ctactactga gatctccagg acagaagtca cctcctctag cagaacatcc  
22081 atctctggtt ctgtgagtc cacaatgttg ccagaaatat ccaccacaag aaaaatcatt  
22141 aagttcccta cttcccccat cctggcagaa tcatcagaaa tgaccatcaa gaccacaaca  
22201 agtctcctg ggtctacatc agagagtacc tttacattag acacatcaac cactccctcc  
22261 ttggtaataa cccattcgac tatgactcag agattgccac actcagagat cagcacctct  
25 22321 gtgagtagag gtgtgggga tgtgccacgg ccagctctc tccctgtgga agaaacaagc  
22381 cctccatctt cccagctgtc tttatctgcc atgatctcac cttctcctgt tcttccaca  
22441 ttaccagcaa gtaggcactc ctcttctgtt tctgtgactt cactctcac accaggccaa  
22501 gtgaagacta ctgaggtgtt ggacgcaagt gcagaacctg aaaccagttc acctccaagt  
22561 ttgagcagca cctcagttga aatactggcc acctctgaag tcaccacaga tacggagaaa  
30 22621 attcatctt tcccaaacac gccagtaacc aaagtggaa cttccagttc tggacatgaa  
22681 tccccctct ctgtctctac tgactcagag acaaccaaag ccacatcggc aatgggtacc  
22741 atctccatta tgggggatac aagtgtttct acattaactc ctgccttacc taacactagg  
22801 aaaattcagt cagagccagc ttcctcactg accaccagat tgaggagagc cagcacctct  
22861 gaagagacca gcttagccac agaagcaaac actgttcttt ctaaagtgtc cactggtgtc  
35 22921 actactgagg tctccaggac agaagccatc tcccttagca gaacatccat gtcaggccct  
22981 gagcagttca caatgtcaca agacatctcc ataggaacca tccccaggat ttctgctctc  
23041 tctgtcctga cagaatctgc aaaaatgacc atcacacccc aaacaggtcc tctggagtct  
23101 aactagaaa gtacccttaa tttgaacaca gcaaccacac cctcttgggt ggaaaccac  
23161 tctatagtaa ttcagggatt tccacaccca gagatgacca cttccatggg cagaggtcct  
40 23221 ggaggtgtgt catggcctag cctccctttt gtgaaagaaa ccagccctcc atcctcccg  
23281 ctgtctttac ctgcccgtgac ctcacctcat cctgtttcca ccacattcct agcacatctc  
23341 cccccctctc ccttctctgt gacttcactt ctcacctctg gcccgcgac aaccacagat  
23401 atcttgggta caagcacaga acctggaacc agttcatctt caagtttgag caccacctcc  
23461 catgagagac tgaccactta caaagacact gcacatacag aagccgtgca tcttccaca  
45 23521 aacacaggag ggaccaatgt ggcaaccacc agctctggat ataaatcaca gtcctctgtc  
23581 ctgactgact catctccaat gtgtaccacc tccacctagg gggatacaag tgttctcaca  
23641 tcaactcctg ccttctctga gactaggagg attcagacag agctagcttc ctcctgacc  
23701 cctggattga gggagtccag tggctctgaa gggaccagct caggcaccac gatgagcaat  
23761 gtctctctta aagtgtccac tgggtctact actgagatct ccaaggaga cgtcacctcc  
50 23821 atcccaggtc ccgctcaatc cacaatatca ccagacatct ccacaagaac cgtcagctgg  
23881 ttctctacat cccctgtcat gacagaatca gcagaaataa ccatgaacac ccatacaagt  
23941 ccttttaggg ccacaacaca aggcaccagt actttggcca cgtcaagcac aacctctttg  
24001 acaatgacac actcaactat atctcaagga ttttcacact cacagatgag cactcttatg  
24061 aggaggggtc ctgaggatgt atcctggatg agccctcccc tcttgaaaaa aactagacct  
55 24121 tctttttctc tgatgtcttc accagccaca acttcacctt ctoctgtttc ctcacattta  
24181 ccagagagca tctcttctc tctcttctct gtgacttca cctctcagtc tggcttggea  
24241 aaaactacag atagtgttca caaaagctca gaacctgtaa ccaactcacc tgcaaatgtg  
24301 agcagcacct cagttgaaat actggccacc tctgaagtca ccacagatac agagaaaaact  
24361 catccttctt caaacagaa agtgaccgat gtggggacct ccagttctgg acatgaatcc  
60 24421 acttctttt tcttagctga ctcacagaca tccaaagtca catctccaat ggttattacc  
24481 tccaccatgg aggatacag tgtctccaca tcaactcctg gcttttttga gactagcaga



	24541	attcagacag	aaccaacatc	ctccctgacc	cttggactga	gaaagaccag	cagctctgag
	24601	gggaccagct	tagccacaga	gatgagcact	gtcctttctg	gagtgccacc	tgggtgccact
	24661	gctgaagtct	ccaggacaga	agtacacctc	tctagcagaa	catccatctc	aggctttgct
	24721	cagctcacag	tgtcaccaga	gacttccaca	gaaaccatca	ccagactccc	tacctccagc
5	24781	ataatgacag	aatcagcaga	aatgatgato	aagacacaaa	cagatocctc	tgggtctaca
	24841	ccagagagta	ctcatactgt	ggacatatca	acaacaccca	actgggtaga	aacccactcg
	24901	actgtgactc	agagattttc	acactcagag	atgaccactc	ttgtgagcag	aagccctggg
	24961	gatatgttat	ggcctagtca	atcctctgtg	gaagaaacca	gctctgcctc	ttccctgctg
	25021	tctctgcctg	ccacgacotc	accttctcct	gtttcctcta	cattagtaga	ggatttccct
10	25081	tccgcttctc	ttcctgtgac	ttctcttctc	acccctggcc	tggtgataac	cacagacagg
	25141	atgggcataa	gcagagaacc	tggaaccagt	tccacttcaa	atgtgagcag	cacctcccat
	25201	gagagactga	ccacttttga	agacactgta	gatacagaag	acatgcagcc	ttccacacac
	25261	acagcagtg	ccaacgtgag	gacctccatt	tctggacatg	aatcacatc	ttctgtccta
	25321	tctgactcag	agacacccaa	agccacatct	ccaatgggta	ccacctacac	catgggggaa
15	25381	acgagtgttt	coatatccac	ttctgacttc	tttgagacca	gcagaattca	gatagaacca
	25441	acatcctccc	tgactctctg	attgagggag	accagcagct	ctgagaggat	cagctcagcc
	25501	acagagggaa	gcactgtcct	ttctgaagtg	cccagtggtg	ctaccactga	ggtctccagg
	25561	acagaagtga	tatcctctag	gggaacatcc	atgtcagggc	ctgatcagtt	cacctatca
	25621	ccagacatct	ctactgaagc	gatcaccagg	ctttctactt	ccccattat	gacagaatca
20	25681	gcagaaagtg	ccatcactat	tgagacaggt	tctcctgggg	ctacatcaga	gggtacctc
	25741	accttggaca	cotcaacaac	aaccttttgg	tcagggaccc	actcaactgc	atctccagga
	25801	ttttcacact	cagagatgac	cactcttatg	agtagaactc	ctggagatgt	gccatggccg
	25861	agccttccct	ctgtggaaga	agccagctct	gtctcttctc	cactgtcttc	acctgccatg
	25921	acctcaactt	cttttttctc	cgattatcca	gagagcatct	cctcctctcc	tcactctgtg
25	25981	actgcacttc	tcacccttgg	cccagtgag	accacagaca	tgttgccgac	aagctcagaa
	26041	cctgaaacca	gttcacctcc	aaatttgagc	agcacctcag	ctgaaatatt	agccacgtct
	26101	gaagtcacca	aagatagaga	gaaaattcat	ccctcctcaa	acacacctgt	agtcaatgta
	26161	gggactgtga	tttataaaca	tctatccctc	tctctgtgtt	tggctgactt	agtgacaaca
	26221	aaacccacat	ctccaatggc	taccacctcc	actctgggga	atacaagtgt	ttccacatca
30	26281	actcctgcct	tcccagaaac	tatgatgaca	cagccaactt	cctccctgac	ttctggatta
	26341	agggagatca	gtacctctca	agagaccagc	tcagcaacag	agagaagtgc	ttctctttct
	26401	ggaatgccca	ctggtgctac	tactaagggt	tccagaacag	aagccctctc	cttaggcaga
	26461	acatccaccc	caggctcctg	tcaatccaca	atatcaccag	aaatctccac	ggaaacacac
	26521	actagaatth	ctactcccct	caccacgaca	ggatcagcag	aaatgaccat	cacccccaaa
35	26581	acaggctcatt	ctggggcatc	ctcacaaggt	acctttacct	tggacacatc	aagcagagcc
	26641	tcttgccag	gaactcactc	agctgcaact	cacagatctc	cacactcagg	gatgaccact
	26701	cctatgagca	gaggctcctga	ggatgtgtca	tggccaagcc	gcccacag	ggaaaaact
	26761	agccctccat	cttccctggg	gtctttatct	gcagtaacct	cacctctgcc	actttattcc
	26821	acaccatctg	agagttagcca	ctcatctcct	ctccgggtga	cttctctttt	cacctctgtc
40	26881	atgatgaaga	ccacagacat	gttgagacaca	agcttgggaa	ctgtgaccac	ttcacctccc
	26941	agtatgaata	tcacctcaga	tgagagtctg	gocacttcta	aagccaccat	ggagacagag
	27001	gcaattcagc	tttcagaaaa	cacagctgtg	actcagatgg	gcacctcag	cgctagacaa
	27061	gaattctatt	cctcttatcc	aggcctccca	gagccatcca	aagtgcacac	tccagtgggtc
	27121	acctcttcca	ccataaaaaga	cattgtttct	acaaccatac	ctgcttctc	tgagataaca
45	27181	agaattgaga	tggagtcaac	atccaccctg	acccccacac	caagggagac	cagcacctcc
	27241	caggagatcc	actcagccac	aaagccaagc	actgttccct	acaaggcact	cactagtgcc
	27301	acgattgagg	actccatgac	acaagtcag	tcctctagca	gaggacctag	ccctgatcag
	27361	tccacaatgt	cacaagacat	atccagtga	gtgatcacca	ggctctctac	ctcccccatc
	27421	aaggcagaat	ctacagaaat	gaccattacc	acccaaacag	gttctcctgg	ggctacatca
50	27481	aggggtaccc	ttaccttgg	cacttcaaca	acttttatgt	cagggaccoc	ctcaactgca
	27541	tctcaaggat	tttcacactc	acagatgacc	gctcttatga	gtagaactcc	tggagatgtg
	27601	ccatggctaa	gccatccctc	tgtggaagaa	gccagctctg	cctctttctc	actgtcttca
	27661	cctgtcatga	cctcatcttc	tccggtttct	tccacattac	cagacagcat	ccactcttct
	27721	tgccttccctg	tgacatcact	tctcacctca	gggctgggtga	agaccacaga	gctgttgggc
55	27781	acaagctcag	aacctgaaac	cagttcaccc	ccaaatttga	gcagcacctc	agctgaaata
	27841	ctggccacca	ctgaagtcac	tacagataca	gagaaactgg	agatgaccaa	tgttgtaacc
	27901	tcaggttata	cacatgaatc	tcttctctct	gtcctagctg	actcagtgac	aacaaaggcc
	27961	acatcttcaa	tgggtatcac	ctacccccca	ggagatacaa	atgttctcac	atcaacccct
	28021	gccttctctg	acaccagtag	gattcaaaaca	aagtcaaagc	tctcactgac	tctcgggttg
60	28081	atggagacca	gcactctctga	agagaccagc	tctgccacag	aaaaaagcac	tgtcctttct
	28141	agtgtgcccc	ctggtgctac	tactgaggtc	tccaggacag	aagccatctc	ttctagcaga

	28201	acatccatcc	caggccctgc	tcaatccaca	atgtcatcag	acacotccat	ggaaccatc
	28261	actagaatth	ctacccccct	cacaaggaaa	gaatcaacag	acatggccat	cacccccaaa
	28321	acaggtcctt	ctggggctac	ctcgaggggt	acctttacct	tggactcatc	aagcacagcc
	28381	tcctggccag	gaactcactc	agctacaact	cagagatttc	cacagtcagt	ggtgacaact
5	28441	cctatgagca	gaggtcctga	ggatgtgtca	tggccaagcc	cgctgtctgt	ggaaaaaaac
	28501	agccctccat	cttccctggg	atcttcatct	tcagtaacct	cacottcgcc	actttattcc
	28561	acaccatctg	ggagtagcca	ctcctctcct	gtccctgtca	cttctotttt	caoctctatc
	28621	atgatgaagg	ccacagacat	gttggatgca	agtttggaa	ctgagaccac	ttcagctccc
	28681	aatatgaata	tcacctcaga	tgagagtctg	gccacttcta	aagccaccac	ggagacagag
10	28741	gcaattcacg	tttttgaaaa	tacagcagcg	tcccatgtgg	aaaccaccag	tgctacagag
	28801	gaactctatt	cctcttcccc	aggcttctca	gagccaacaa	aagtgatata	tccagtggtc
	28861	acctcttctc	ctataagaga	caacatgggt	tccacaacaa	tgctgggtc	ctctggcatt
	28921	acaaggattg	agatagagtc	aatgtcatct	ctgacccctg	gactgaggga	gaccagaacc
	28981	tcccaggaca	tcacctcatc	cacagagaca	agcactgtcc	tttacaagat	gtcctctggg
15	29041	gccactcctg	agggtctccag	gacagaagtt	atgccctcta	gcagaacatc	cattcctggc
	29101	cctgctcagt	ccacaatgtc	actagacatc	tccgatgaag	ttgtcaccag	gctgtctacc
	29161	tctcccatca	tgacagaatc	tgacagaaata	accatcacca	cccaaacagg	ttattctctg
	29221	gctacatccc	aggttaccct	tcccttgggc	acctcaatga	cctttttgtc	agggaccacc
	29281	tcaactatgt	ctcaaggact	ttcacactca	gagatgacca	atcttatgag	caggggtcct
20	29341	gaaagtctgt	catggacgag	cctctgcttt	gtggaaacaa	ctagatcttc	ctcttctctg
	29401	acatcattac	ctctcacgac	ctcactttct	cctgtgtcct	ccacattact	agacagtagc
	29461	ccctcctctc	ctcttctctg	gacttcactt	atcctcccag	gctgggtgaa	gactacagaa
	29521	gtgttgata	caagctcaga	gcctaaaacc	agttcatctc	caaatttgag	cagcacctca
	29581	gttgaaatac	cggccacctc	tgaatcatg	acagatacag	agaaaattca	tccttctca
25	29641	aacacagcgg	tggccaaagt	gaggacctcc	agttctgttc	atgaatctca	ttcctctgtc
	29701	ctagctgact	cagaaacaac	cataaaccata	ccttcaatgg	gtatcacctc	cgctgtggac
	29761	gataccactg	ttttcacatc	aaatcctgcc	ttctctgaga	ctaggaggat	tccgacagag
	29821	ccaacattct	cattgactcc	tggattcagg	gagactagca	cctctgaaga	gaccacctca
	29881	atcacagaaa	caagtgcagt	cctttatgga	gtgcccacta	gtgctactac	tgaagtctcc
30	29941	atgacagaaa	tcatgtcctc	taatagaaca	cacatccctg	actctgatca	gtccacgatg
	30001	tctccagaca	tcactactga	agtgtatcac	aggctctctt	cctcatccat	gatgtcagaa
	30061	tcaacacaaa	tgaccatcac	cacccaaaaa	agttctcctg	gggctacagc	acagagtact
	30121	cttaccttgg	ccacaacaac	agcccccttg	gcaaggaccc	actcaactgt	tcctcctaga
	30181	tttttacct	cagagatgac	aactcttatg	agtaggagtc	ctgaaaatcc	atcatggaag
35	30241	agctctccct	ttgtgaaaaa	aaactagctc	tcactttctc	tgttgtcctt	acctgtcacg
	30301	acctcacctt	ctgtttcttc	cacattaccg	cagagtatcc	cttctctctc	tttttctgtg
	30361	acttcaactc	tcaccccagg	catgggtgaag	actacagaca	caagcacaga	acctggaacc
	30421	agtttatctc	caaattctgag	tggcacctca	gttgaataac	tggctgcctc	tgaagtccac
	30481	acagatacag	agaaaattca	tccttcttca	agcatggcag	tgaccaatgt	gggaaccacc
40	30541	agttctggac	atgaactata	ttcctctgtt	tcaatccact	cggagccatc	caaggctaca
	30601	tacccagtgg	gtactccctc	ttccatgggt	gaaacctcta	tttccacatc	aatgctgtct
	30661	aattttgaga	ccacaggatt	tgaggctgag	ccattttctc	atttgacttc	tggatttagg
	30721	aagacaaaac	tgtccctgga	caccagctca	gtcacaccaa	caaatacacc	ttcttctcct
	30781	gggtccactc	accttttaca	gagttccaag	actgatttca	cctcttctgc	aaaaacatca
45	30841	tccccagact	ggcctccagc	ctcacagtat	actgaaattc	cagtggacat	aatcaccccc
	30901	tttaatgctt	ctccatctat	tacggagtcc	actgggataa	cctccttccc	agaaatccagg
	30961	tttactatgt	ctgtaacaga	aagtactcat	catctgagta	cagatttgct	gccttcagct
	31021	gagactatth	ccactggcac	agtgtatgct	tctctatcag	aggccatgac	ttcatttgcc
	31081	accactggag	ttccacgagc	catctcaggt	tcaggtagtc	cattctctag	gacagagtca
50	31141	ggccctgggg	atgctactct	gtccaccatt	gcagagagcc	tgccctcatc	cactcctgtg
	31201	ccattctcct	cttcaacctt	cactaccact	gattcttcaa	ccatcccagc	cctccatgag
	31261	ataacttcc	cttccagctac	cccatataga	gtggacacca	gtcttgggac	agagagcagc
	31321	actactgaag	gacgcttggg	tatgggtcagt	actttggaca	cttcaagcca	accaggcagg
	31381	acatcttcaa	cacccatttt	ggataccaga	atgacagaga	gcgttgagct	gggaacagtg
55	31441	acaagtgtct	atcaagttcc	ttcactctca	acacggttga	caagaactga	tggcattatg
	31501	gaacacatca	caaaaatacc	caatgaagca	gcacacagag	gtaccataag	accagtcaaa
	31561	ggccctcaga	catccacttc	gctgtccagt	cctaaaggac	tacacacagg	agggacaaaa
	31621	agaatggaga	ccaccaccac	agctttgaag	accaccacca	cagctttgaa	gaccacttcc
	31681	agagccacct	tgaccaccag	tgtctatact	cccactttgg	gaacactgac	tcccctcaat
60	31741	gcatcaaggc	aaatggccag	cacaatcctc	acagaaatga	tgatcacaa	cccataatgt
	31801	ttccctgatg	ttccagaaac	gacatcctca	ttggctacca	gcctgggagc	agaaaccagc

	31861	acagctcttc	ccaggacaac	cccattctgtt	ctcaatagag	aatcagagac	cacagcctca
	31921	ctggtctctc	gttctggggc	agagagaagt	cgggttattc	aaactctaga	tgtttcttct
	31981	agtgagccag	atacaacagc	ttcatgggtt	atccatctcg	cagagaccat	cccaactggt
	32041	tccaagacaa	cccccaattt	ttccacagt	gaattagaca	ctgtatcttc	cacagccacc
5	32101	agtcattggg	cagacgtcag	ctcagccatt	ccaacaaata	tctcacctag	tgaactagat
	32161	gcaactgacc	cactgggtcac	tatttcgggg	acagatacta	gtacaacatt	cccaactggt
	32221	actaagtcac	cacatgaaac	agagacaaga	accacatggc	tcactcatcc	tgcagagacc
	32281	agctcaacta	ttcccagaac	aatccccaat	ttttctctac	atgaatcaga	tgccacacct
	32341	tcaatagcca	ccagtctctg	ggcagaaacc	agttcagcta	ttccaattat	gactgtctca
10	32401	cctgggtgag	aagatctggt	gacctcagag	gtcactagtt	ctgggacaga	cagaaatagc
	32461	actattccaa	ctttgactct	ttctcctggt	gaaccaaaga	cgatagcctc	attagtcacc
	32521	catcctgaag	cacagacaag	ttcggccatt	ccaacttcaa	ctatctcgcc	tgctgtatca
	32581	cgggttggtg	cctcaatggt	caccagtttg	gcggcaaaga	caagtacaac	taatcgagct
	32641	ctgacaaact	cccctgggtg	accagctaca	acagtttcat	tggtcacgca	tcctgcacag
15	32701	accagcccaa	cagttccctg	gacaaacttc	atTTTTttcc	atagtaaatc	agacaccaca
	32761	ccttcaatga	ccaccagtca	tggggcagaa	tccagttcag	ctgtttccaa	ttccaactgtt
	32821	tcaactgagg	taccaggagt	agtgaacctc	ttggtcacca	gttctagggc	agtgatcagt
	32881	acaactattc	caattctgac	tccttctcct	ggtgaaaccg	agaccacacc	ttcaatggcc
	32941	accagtcag	gggaagaagc	cagttctgct	attccaactc	caactgtttc	acctggggta
20	33001	ccaggagtgg	tgacctctct	gttcaactag	tctagggcag	tgactagtac	aactattcca
	33061	attctgactt	tttctcttgg	tgaaccagag	accacacctt	caatggccac	cagtcattggg
	33121	acagaagctg	gctcagctgt	ttccaactgtt	ttacctgagg	taccaggagt	ggtagctctt
	33181	ctggttgcta	gttctagggc	agtaaccagt	acaactcttc	caactctgac	tccttctcct
	33241	ggtagaacag	agaccacacc	ttcaatggcc	accagtcag	gggcagaagc	cagctcaact
25	33301	gttccaactg	tttccacctg	ggtaccagga	gtggtgacct	ctctgggtcac	tagttctagt
	33361	ggagtaaaac	gtacaagtat	ttccaactctg	attctttctc	ctggtgaact	agaaaccaca
	33421	ccttcaatgg	ccaccagtca	tggggcagaa	gccagctcag	ctgttccaa	ttccaactgtt
	33481	tcacctgggg	tatcaggagt	ggtgacctct	ctggtcacta	gttccagggc	agtgaccagt
	33541	acaactattc	caattctaac	tccttcttct	agtgagccag	agaccacacc	ttcaatggcc
30	33601	accagtcag	gggtagaagc	cagctcagct	gttctaactg	tttccacctg	ggtaccagga
	33661	atgggtgacct	ctctggtcac	tagttctaga	gcagtaacca	gtacaactat	ttccaactctg
	33721	actatttctt	ctgatgaacc	agagaccaca	acttcattgg	tcacctattc	tgaggcaaat
	33781	atgatttcag	ccatttccaa	tttagctgtc	ttccctactg	tacaagggtc	ggtgacttca
	33841	ctggtcacta	gttctgggtc	agagaccagt	gcgttttcaa	atctaactgt	tgcttcaagt
35	33901	caaccagaga	ccatagactc	atgggtcgtc	catcctggga	cagaagcaag	ttctgttggt
	33961	ccaactttga	ctgtctccac	tggtgagccg	tttcaaaata	ttctattggt	cacctatcct
	34021	gcagagagta	gctcaactct	ttccaggaca	acctcaaggt	tttcccacag	tgaattagac
	34081	actatgcctt	ctacagtcac	cagtcctgag	gcagaatcca	gctcagccat	ttcaactact
	34141	attttcacctg	gtataccagg	tggtgctgac	tcactggtca	ctagctctgg	gagagacatc
40	34201	agtgaactct	ttccaacagt	gctgagtcac	ccacatgaat	cagaggccaac	agcctcatgg
	34261	gttactcatc	ctgcagtcac	cagcacaaca	gttcccaggga	caaccctcaa	ttattctcat
	34321	agtgaaccag	acaccacacc	atcaatagcc	accagtcctg	gggcagaagc	caacttcagt
	34381	tttccaacaa	taactgtctc	acctgatgta	ccagatatgg	taacctcaca	ggtcactagt
	34441	ctctgggacag	acaccagtat	aactattcca	actctgactc	tttcttctgg	tgagccagag
45	34501	accacaacct	catttatcac	ctattctgag	acacacacaa	gttcagccat	ttcaactctc
	34561	cctgtctccc	ctggtgcatc	aaagatgctg	acctcactgg	tcactcagtt	tgggcagac
	34621	agcactacaa	ctttcccaac	actgacggag	accccatatg	aaccagagac	aacagccata
	34681	cagctcattc	atcctgcaga	gaccaacaca	atgggtccca	agacaactcc	caagttttcc
	34741	catagtaagt	cagacaccac	actcccagta	gccatcacca	gtcctggggc	agaagccagt
50	34801	tcagctgttt	caacgacaac	tatctcacct	gatagtgcag	atctggtgac	ctcactgggtc
	34861	cctagttctg	ggacagacac	cagtacaacc	ttcccaacat	tgagtgcagc	cccatatgaa
	34921	ccagagacta	cagtcacgtg	gctcactcat	cctgcagaaa	ccagcacaac	ggtttctggg
	34981	acaattccca	acttttccca	taggggatca	gacactgcac	cctcaatggt	caccagtcct
	35041	ggagtagaca	cgagggtcag	tggttccaact	acaacctacc	caccagtagt	accaggggta
55	35101	gtgacctcac	aggtcactag	ttctgcaaca	gacactagta	cagctatttc	aacttttgat
	35161	ccttctcctg	gtgaaccaga	gaccacagcc	tcactcagct	cccatcctgg	gacacagact
	35221	ggcttcactg	ttccaattcg	gactgttccc	tcagtgcagc	cagatcaaat	ggcttctcgg
	35281	gtcactcatc	ctccacagac	cagcacacct	gtttccagaa	caacctccag	tttttcccat
	35341	agtagtcocg	tgtaatggcc	accagtccta	ggacagaagc	cagttcagct	cactagtctc
60	35401	gtactgacaa	caatctcacc	tggtgcacca	gagatggtga	cttcacagat	cactagtctc
	35461	ggggcagcaa	ccagtcacaac	tggttccaact	ttgactcatt	ctcctgggtat	gccagagacc

	35521	acagccttat	tgagcaccca	tcccagaaca	gggacaagta	aaacatttcc	tgcttcaact
	35581	gtgtttcctc	aagtatcaga	gaccacagcc	tcactcacca	ttagacctgg	tgacagagact
	35641	agcacagctc	tcccactca	gacaacatcc	tctctcttca	ccctacttgt	aactggaacc
	35701	agcagagtgg	atctaagtcc	aactgcttca	cctggtgttt	ctgcaaaaac	agccccactt
5	35761	tccacccatc	cagggacaga	gaccagcaca	atgattccaa	cttcaactct	ttcccttggt
	35821	ttactagaga	ctacaggctt	actggccacc	agctcttcag	cagagaccag	cacgagtact
	35881	ctaactctga	ctgtttcccc	tgtgtctct	gggctttcca	gtgcctctat	aacaactgat
	35941	aagcccaaaa	ctgtgacctc	ctggaacaca	gaaacctcac	catctgtaac	ttcagttgga
	36001	cccccagaat	tttccaggac	tgtcacaggc	accactatga	ccttgatacc	atcagagatg
10	36061	ccaacaccac	ctaaaaccag	tcattggaga	ggagtgaagc	caaccactat	cttgagaact
	36121	acaatggttg	aagccactaa	tttagctacc	acaggttcca	gtccactgt	ggccaagaca
	36181	acaaccacct	tcaatacact	ggctggaagc	ctctttactc	ctctgaccac	acctgggatg
	36241	tccaccttgg	cctctgagag	tgtgacctca	agaacaagtt	ataaccatcg	gtcctggatc
	36301	tccaccacca	gcagttataa	cgtctggtag	tggacccctg	ccaccagcac	tccagtgaat
15	36361	tctacattct	ccccaggat	ttccacatcc	tccatcccca	gctccacagc	agccacagtc
	36421	ccattcatgg	tgccattcac	cctcaacttc	accatcacca	acctgagta	cgaggaggac
	36481	atgctggcacc	ctggttccag	gaagtccaac	gccacagaga	gagaactgca	gggtctgctc
	36541	aaacctctgt	tcaggaatag	cagtctggaa	tacctctatt	caggctgcag	actagcctca
	36601	ctcaggccag	agaaggatag	ctcagccatg	gcagtggatg	ccatctgcac	acatgcacct
20	36661	gacctgaag	acctcgact	ggacagagag	cgactgtact	gggagctgag	caatctgaca
	36721	aatggcatcc	aggagctggg	cccctacacc	ctggaccgga	acagtctcta	tgtcaatggt
	36781	ttcaccatcc	gaagctctat	gcccaccacc	agcactcctg	ggacctccac	agtggatgtg
	36841	ggaacctcag	ggactccatc	ctccagcccc	agccccacgg	ctgctggccc	tctctgtatg
	36901	ccgtttaccc	tcaacttcac	catcaccaac	ctgcagtacg	aggaggacat	gcgtgcacat
25	36961	ggctccagga	agttcaacac	catggagagt	gtcctgcagg	gtctgtctaa	gcccttggtc
	37021	aagaacacca	gtgttggccc	tctgtactct	ggctgcagat	tgaccttgtc	caggcccgag
	37081	aaagatgggg	cagccactgg	agtggatgcc	atctgcaccc	accgccttga	ccccaaaagc
	37141	cctggactca	acagggagca	gctgtactgg	gagctaagca	aactgacca	tgacattgaa
	37201	gagctggggc	cctacacct	ggacaggaac	agtctctatg	tcaatgggtt	cacctcatcg
30	37261	agctctgtgt	ccaccaccag	cactcctggg	acctccacag	tggatctcag	aaacctcagg
	37321	actccatcct	ccctctccag	ccccacaatt	atggctgtct	gccctctcct	ggtaccattc
	37381	accttcaact	tcaccatcac	caacctgcag	tatggggagg	acatgggtca	ccctggctcc
	37441	agggaagtta	acaccacaga	gagggtcctg	cagggtctgc	ttggtcccat	attcaagaac
	37501	accagtgttg	gccctctgta	ctctggctgc	agactgacct	ctctcaggtc	tgagaaggat
35	37561	ggagcagcca	ctggagtggg	tgccatctgc	atccatcatc	ttgaccccaa	aagccctgga
	37621	ctcaacagag	agcggctgta	ctgggagctg	agccaactga	ccaatggcat	caaagagctg
	37681	ggcccctaca	ccctggacag	gaacagtctc	tatgtcaatg	gtttcaccca	ctggacctct
	37741	gtgcccacca	ccagcactcc	tgggacctcc	acagtggacc	ttggaaacct	agggactcca
	37801	ttctccctcc	caagccccgc	aactgctggc	cctctcctgg	tgtgttcac	cctcaacttc
40	37861	accatcacca	acctgaagta	tgaggaggac	atgcatcgcc	ctggctccag	gaagtccaac
	37921	accactgaga	gggtcctgca	gactctgctt	ggtcctatgt	tcaagaacac	cagtgttggt
	37981	cttctgtact	ctggctgcag	actgaccttg	ctcagggtcc	agaaggatgg	agcagccact
	38041	ggagtggatg	ccatctgcac	ccacctcttt	gaccccaaaa	gccctggact	ggacagagag
	38101	cagctatact	gggagctgag	ccagctgacc	aatggcatca	aagagctggg	cccctacacc
45	38161	ctggacagga	acagtctcta	tgtcaatggt	ttcacccatt	ggatccctgt	gcccaccagc
	38221	agcactcctg	ggacctccac	agtggacctt	gggtcaggga	ctccatcctc	cctccccage
	38281	cccacagctg	ctggccctct	cctggtgcca	ttcacccctc	acttcacat	caccaacctg
	38341	cagtacgagg	aggacatgca	tcaccagggc	tccaggaagt	tcaacaccac	ggagcgggtc
	38401	ctgcaggggc	tgcttgggtc	catgttcaag	aacaccagtg	tcggccttct	gtactctggc
50	38461	tgacagactga	ccttgctcag	gtccgagaag	gatggagcag	ccactggagt	ggatgccatc
	38521	tgacccacc	gtcttgacct	caaaagccct	ggagtggaca	gggagcagct	atactgggag
	38581	ctgagccagc	tgaccaatgg	catcaaagag	ctgggtccct	acacctgga	cagaacagct
	38641	ctctatgtca	atggtttcac	ccatcagacc	tctgcgcccc	acaccagcac	tcctgggacc
	38701	tccacagtgg	accttgggac	ctcagggact	ccatcctccc	tccccagccc	tacatcngct
55	38761	ggccctctcc	tggtnccttt	cacctccaac	ttcacctca	ccaacctgca	gtacgaggag
	38821	gacatgcggc	accnngntc	caggaaagttc	aacaccacng	agagggtnct	gcagggtctg
	38881	ctnaagcccc	tnttcaagag	caccagtgtt	ggccctctgt	actctggctg	cagactgacc
	38941	ttgctcaggt	cogagaagga	tggagcagcc	actggagtgg	atgccatctg	caccacacct
	39001	cttgacccca	aaagccctgg	agtggacagg	gagcagctat	actgggagct	gagcagctg
60	39061	accaatggca	tcaaagagct	gggtccctac	acctgggaca	gaaacagctc	ctatgtcaat
	39121	ggtttcaccc	atcagacctc	tgccgccaac	accagcactc	ctgggacctc	cacagtggac

39181 cttgggacct cagggaactoc atcctccctc cccagcccta catctgctgg ccctctcctg  
 39241 gtgccattca ccctcaactt caccatcacc aacctgcagt acgaggagga catgcatcac  
 39301 ccaggctcca ggaagttcaa caccacggag cgggtcctgc agggctctgt tgggtcccatg  
 39361 ttcaagaaca ccagtgtcgg ccttctgtac tctggctgca gactgacctt gctcaggcct  
 5 39421 gagaagaatg gggcagccac tgggaatggat gccatctgca gccaccgtct tgaccccaaa  
 39481 agccctggac tcaacagaga gcagctgtac tgggagctga gccagctgac ccattggcatc  
 39541 aaagagctgg gcccctacac cctggacagg aacagtctct atgtcaatgg tttcaccat  
 39601 cggagctctg tggcccccac cagcactcct gggacctcca cagtggacct tgggacctca  
 39661 gggactccat cctccctccc cagcccccaca acagctgttc ctctcctggt gccgttcacc  
 10 39721 ctcaacttta ccattaccaa tctgcagtat ggggaggaca tgcgtcacc ccgtccagg  
 39781 aagttcaaca ccacagagag ggtcctgcag ggtctgcttg gtcccttgtt caagaactcc  
 39841 agtgtcggcc ctctgtactc tggctgcaga ctgatctctc tcagggtctga gaaggatggg  
 39901 gcagccactg gagtggatgc catctgcacc caccacctta accctcaaag ccttggactg  
 39961 gacagggagc agctgtactg gcagctgagc cagatgacca atggcatcaa agagctgggc  
 15 40021 ccctacaccc tggaccggaa cagtctctac gtcaatggtt tccccatcg gagctctggg  
 40081 ctcaccacca gcactccttg gacttccaca gttgaccttg gaacctcagg gactccatcc  
 40141 cccgtcccca gcccacaaac tgetggccct ctctggtg cttcaccct caacttcacc  
 40201 atcaccaacc tgcagtatga ggaggacatg catcgccctg gatctaggaa gttcaacacc  
 20 40261 acagagaggg tctgcaggg tctgcttagt cccattttca agaactccag tgggtggccct  
 40321 ctgtactctg gctgcagact gacctctctc agggcccgaga aggatggggc gactccatcc  
 40381 atggatgctg tctgcctcta ccacctaat cccaaaagac ctggactgga cagagagcag  
 40441 ctgtactggg agctaagcca gctgaccac aacatcactg agctggggcc ctacagcctg  
 40501 gacagggaca gtctctatgt caatggtttc acccatcaga actctgtgct caccaccagt  
 40561 actcctggga cctccacagt gtactgggca accactggga ctccatcctc cttccccggc  
 25 40621 cacacagagc ctggccctct cctgatacca ttcaatttca actttaccat caccacactg  
 40681 cattatgagg aaaacatgca acaccctggt tccagggaagt tcaacaccac ggagagggtt  
 40741 ctgcaggggtc tgctcaagcc ctgtgtcaag aacaccagtg ttggccctct gtactctggc  
 40801 tgcagactga cctctctcag gcccgagaag gatggggcag caactggaat ggtatgctgc  
 30 40861 tgcctctacc accctaatac caaaagacct gggctggaca gagagcagct gtactgggag  
 40921 ctaagccagc tgaccacaa catcactgag ctgggcccct acagcctgga caggagacagt  
 40981 ctctatgtca atggtttcac ccatacagaac tctgtgccca ccaccagtac tctgggacc  
 41041 tccacagtgt actgggcaac cactgggact ccatcctcct tccccggcca cacagagcct  
 41101 gggccctctc tgataccatt cactttcaac tttaccatca ccaacctgca ttatgaggaa  
 35 41161 aacatgcaac accctggttc cagggaagttc aacaccacgg agagggttct gcagggtctg  
 41221 ctcaagccct tgttcaagaa caccagtgtt ggccctctgt actctggctg cagactgacc  
 41281 ttgctcagac ctgagaagca tgaggcagcc actggagtggt acaccatctg taccaccgc  
 41341 gttgatccca tgcgacctgg actggacagg gagcggtat actgggagct gagccagctg  
 41401 accaacagca ttaccgaact gggacctac accctggaca gggacagctc ctatgtcaat  
 41461 ggttcaacc ctgcgagctc tgtgccaacc accagcactc ctgggacctc cacagtgcac  
 40 41521 ctggcaacct ctgggactcc atcctccctg cctggccaca cagccctgt cctctctctg  
 41581 ataccattca ccctcaactt taccatcacc aacctgcatt atgaggaaaa catgcaacac  
 41641 cctggttcca ggaagttcaa caccacggag aggggtctgc agggctctgt caagccctg  
 41701 ttcaagaaca ccagtgttg ccctctgtac tctggctgca gactgacctt gctcagacct  
 41761 gagaagcatg aggcagccac tggagtgga accatctgta cccaccgctg tgatcccatc  
 45 41821 ggacctggac tgnacagnga gcngctntac tgggagctna gccanctgac caannncatc  
 41881 nnnagctgg gmcctacac cctggacagg nacagtctct atgtcaatgg tttcaccat  
 41941 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn  
 42001 gggactccat cctcctccc cngcncaca tctgctggcc ctctcctggt gccattcacc  
 42061 ctcaacttca ccattaccaa cctgcagtac gaggaggaca tgcattcacc aggtccagg  
 50 42121 aagttcaaca ccacggagcg ggtcctgcag ggtctgcttg gtcccatgtt caagaacacc  
 42181 agtgtcggcc ttctgtactc tggctgcaga ctgaccttgc tcaggcctga gaagaatggg  
 42241 gcagccactg gaattggatgc catctgcagc cacogtcttg accccaaaag ccttggactc  
 42301 gacagagagc agctgtactg ggagctgagc cagctgacct atggcatcaa agagctgggc  
 42361 ccctacaccc tggacaggaa cagtctctat gtcaatgggt tcacctatcg gagctctgtg  
 55 42421 gcccaccca cactcctgg gactccaca gtggaccttg ggacctcagg gactccatcc  
 42481 tccctcccca gcccacaaac agctgttctt ctctggtg cgttaccct caactttacc  
 42541 atcaccaatc tgcagtatgg ggaggacatg cgtcacctg gctccaggaa gttcaacacc  
 42601 acagagaggg tctgcaggg tctgcttggt ccttgttca agaactccag tgtcggccct  
 42661 ctgtactctg gctgcagact gatctctctc aggtctgaga aggatggggc agccactgga  
 60 42721 gtggatgcca tctgcaccca ccacctaac cctcaaagcc ctggactgga caggagcag  
 42781 ctgtactggc agctgagcca gatgaccaat ggcatcaaag agctggggcc ctacacctg

	42841	gaccggaaca	gtctctacgt	caatgggttc	acccatcgga	gctctgggct	caccaccagc
	42901	actccttgga	cttccacagt	tgaccttgga	acctcagggg	ctccatcccc	cgtccccagc
	42961	cccacaactg	ctggccctct	cctgggtgcca	ttcacccctaa	acttcacccat	caccaacctg
	43021	cagtattgagg	aggacatgca	tcgcccctgga	tctaggaagt	tcaacgccac	agagagggtc
5	43081	ctgcagggtc	tgcttagtcc	catattcaag	aactccagtg	ttggccctct	gtactctggc
	43141	tgccagactga	cctctctcag	gcccggagaag	gatggggcag	caactggaat	ggatgctgtc
	43201	tgccctctacc	accctaatcc	caaaagacct	ggactggaca	gagagcagct	gtactgggag
	43261	ctaagccagc	tgaccacaa	catcactgag	ctgggcccct	acagcctgga	caggggacagt
	43321	ctctatgtca	atgggtttcac	ccatcagagc	tctatgacga	ccaccagaac	tcctgatacc
10	43381	tcacacaatgc	acctggcaac	ctcgagaact	ccagcctccc	tgtctggacc	tacgaccgcc
	43441	agccctctcc	tggtgctatt	cacaatcaac	tgacccatca	ccaacctgca	gtacgaggag
	43501	gacatgctc	gcactggctc	caggaagtcc	aacacccatgg	agagtgtcct	gcagggtctg
	43561	ctcaagccct	tggttcaagaa	caccagtgtt	ggccctctgt	actctggctg	cagattgacc
	43621	ttgctcaggc	ccaagaaaga	tggggcagcc	actggagtg	atgccatctg	caccaccgcc
15	43681	cttgacccca	aaagccctgg	actcaacagg	gagcagctgt	actgggagct	aagcaaaactg
	43741	accaatgaca	ttgaagagct	gggcccctac	acctgggaca	ggaacagtct	ctatgtcaat
	43801	ggtttcaccc	atcagagctc	tgtgtccacc	accagcactc	ctgggacctc	cacagtggat
	43861	ctcagaacct	cagggactcc	atcctccctc	tccagcccca	caattatgnc	nnctgnccct
	43921	ctcctgntnc	cnttccacct	caacttnacc	atcaccaacc	tgcantangn	ggannacatg
20	43981	cnncnccng	gntccaggaa	gttcaacacc	acngagagg	tcctacaggg	tcctgctcag
	44041	cccttggtca	agaacaccag	tgtagctct	ctgtactctg	gttgccagact	gacctgtctc
	44101	aggcctgaga	aggatggggc	agccaccaga	gtggatgctg	cctgcaccta	ccgcccctgat
	44161	cccaaaagcc	ctggactgga	cagagagcaa	ctatactggg	agctgagcca	gctaaccac
	44221	agcatcactg	agctgggacc	ctacaccctg	gacaggggtca	gtctctatgt	caatggcttc
25	44281	aacctctgga	gctctgtgcc	aaccaccagc	actcctggga	cctccacagt	gcactgggca
	44341	acctctggga	ctccatcctc	cctgcctggc	cacacancnn	ctgnccctct	cctgntnccn
	44401	ttcacentca	acttnaccat	caccaacctg	cantangngg	annacatgcn	nncccnnggn
	44461	tcagggaagt	tcaacaccac	ngagagggtt	ctgcagggtc	tgctcaaacc	cttgttcagg
	44521	aatagcagtc	tggaataacct	ctattcaggc	tgccagactag	cctcactcag	gccagagaag
30	44581	gatagctcag	ccatggcagt	ggatgccatc	tgccacacatc	gccctgaccc	tgaagacctc
	44641	ggactggaca	gagagcgact	gtactgggag	ctgagcaatc	tgcaaatgg	catccaggag
	44701	ctggggccct	acaccctgga	ccggaacagt	ctctacgtca	atggtttcac	ccatcggagc
	44761	tctgggctca	ccaccagcac	tccttggaact	tccacagttg	accttggaac	ctcagggaact
	44821	ccatcccccg	tccccagccc	cacaactgct	ggccctctcc	tggtgccatt	cacctcaac
35	44881	ttcacccatca	ccaacctgca	gtatgaggag	gacatgcac	gccctgggtc	caggagggtc
	44941	aacaccacgg	agagggttct	gcagggtctg	ctcagccct	tggtcaagaa	caccagtgtt
	45001	ggccctctgt	actctggctg	cagactgacc	tgctcagac	ctgagaagca	agaggcagcc
	45061	actggagtg	acaccatctg	taccacccgc	gttgatccca	tcggacctgg	actggacaga
	45121	gagcggctat	actgggagct	gagccagctg	accaacagca	tcacagagct	gggacctac
40	45181	acctgggata	gggacagtct	ctatgtcaat	ggcttcaacc	cttgaggctc	tggtccaacc
	45241	accagcactc	ctgggacctc	cacagtgcac	ctggcaacct	ctgggactcc	atcctccctg
	45301	cctggccaca	cagccctgtg	ccctctcttg	ataccattca	ccctcaactt	taccatcacc
	45361	gacctgcatt	atgaagaaaa	catgcaaac	cctgggtcca	ggaagttcaa	caccacggag
	45421	aggggttctgc	agggctctgt	caagcccttg	ttcaagagca	ccagcgttg	ccctctgtac
45	45481	tctggctgca	gactgacctt	gctcagacct	gagaaacatg	gggcagccac	tggagtggac
	45541	gccatctgca	ccctccgctt	tgatcccaact	ggctcctggac	tggaacagaga	gcggctatac
	45601	tgggagctga	gccagctgac	caacagcgtt	acagagctgg	gcccctacac	cctggacagg
	45661	gacagtctct	atgtcaatgg	cttcacccat	cggagctctg	tgccaaccac	cagtattcct
	45721	gggacctctg	cagtgcacct	ggaaacctct	gggactccag	cctccctccc	tgccacacac
50	45781	gccccctggcc	ctctcctggt	gcatttcacc	ctcaacttca	ctatcaccaa	cctgcagtat
	45841	gaggaggaca	tgctgcaccc	tggttcagg	aagttcagca	ccacggagag	agtccctgag
	45901	ggtctgctca	agcccttgtt	caagaacacc	agtgtcagct	ctctgtactc	tggttgacga
	45961	ctgaccttgc	tcaggcctga	gaaggatggg	gcagccacca	gagtggatgc	tgctgtcacc
	46021	catcgtcctg	accccaaaag	ccctggactg	gacagagagc	ggctgtactg	gaagctgagc
55	46081	cagctgaccc	acggcatcac	tgagctgggc	ccctacaccc	tggaacaggca	cagctctcat
	46141	gtcaatgggtt	tcacccatca	gagctctatg	acgaccacca	gaactcctga	tacctccaca
	46201	atgcacctgg	caacctcgag	aactccagcc	tcctgtgtctg	gacctacgac	cgccagccct
	46261	ctcctgggtg	tattcacaat	taacttcacc	atcactaac	tgccgtatga	ggagaacatg
	46321	catcacccctg	gctctagaaa	gtttaacacc	acggagagag	tccttcagg	tctgtcagg
60	46381	cctgtgttca	agaacaccag	tggtggccct	ctgtactctg	gctgcagact	gaccacgctc
	46441	agggccaaga	aggatggggc	agccacccaa	gtggatgcca	tctgcaccta	ccgcccctgat

	46501	cccaaaagcc	ctggactgga	cagagagcag	ctatactggg	agctgagcca	gctaaccacc
	46561	agcatcactg	agctggggccc	ctacaccacg	gacaggagaca	gtctctatgt	caatggcttc
	46621	acccatcggg	gctctgtgcc	aaccaccagt	attcctggga	cctctgcagt	gcacctggaa
	46681	acctctggga	ctccagcctc	cctccctggc	cacacagccc	ctggccctct	cctgggtgcca
5	46741	ttcaccctca	acttcaactat	caccaacctg	cagtatgagg	aggacatgcg	tcaccctggt
	46801	tccaggaagt	tcaacaccac	ggagagagtc	ctgcagggtc	tgctcaagcc	cttgttcaag
	46861	agcaccagt	ttggccctct	gtactctggc	tgcagactga	ccttgctcag	gcctgaaaaa
	46921	cgtggggcag	ccaccggcgt	ggacaccatc	tgcactcacc	gccttgaccc	tctaaccaca
	46981	ggactggaca	gagagcagct	atactgggag	ctgagcaaac	tgaccctggt	catcatcgag
10	47041	ctggggccct	acctcctgga	cagaggcagt	ctctatgtca	atggtttcac	ccatcgagcc
	47101	tctgtgcccc	ccaccagcac	tcctggggacc	tccacagtgg	accttggaac	ctcagggact
	47161	ccattctccc	tcccaagccc	cgancnnct	gnccctctcc	tgntnccntt	caccntcaac
	47221	ttnacatca	ccaacctgca	ntangnggan	nacatgcnnc	ncccnngntc	cagggaagttc
	47281	aacaccacng	agagggtcct	gcagactctg	cttggctcta	tgttcaagaa	caccagtgtt
15	47341	ggccttctgt	actctggctg	cagactgacc	ttgctcaggt	ccgagaagga	tggaagcacc
	47401	actggagtgg	atgccatctg	cacccaccgt	cttgacccca	aaagccctgg	agtggacagg
	47461	gagcaactat	actgggagct	gagccagctg	accaatggca	ttaaagaact	gggccccctac
	47521	acctgggaca	ggaacagtct	ctatgtcaat	gggttcaccc	attggatccc	tgtgcccacc
	47581	agcagcactc	ctgggacctc	cacagtggac	cttgggtcag	ggactccatc	ctccctcccc
20	47641	agccccacaa	ctgctggccc	tctcctgggt	ccgttcaccc	tcaacttcac	catcaccaac
	47701	ctgaagtacg	aggaggacat	gcattgccct	ggctccagga	agttaacac	cacagagaga
	47761	gtcctgcaga	gtctgcttgg	tcccatgttc	aagaacacca	gtgttggtcc	tctgtactct
	47821	ggctgcagac	tgaccttgc	caggtccgag	aaggatggag	cagccactgg	agtggatgcc
	47881	atctgcaccc	acogtcttga	cccacaaagc	cctggagtgg	acagggagca	gctatactgg
25	47941	gagctgagcc	agctgaccaa	tggcatcaaa	gagctgggtc	cctacacctc	ggacagaaac
	48001	agtctctatg	tcaatggttt	caccatcag	acctctgcgc	ccaacaccag	cactcctggg
	48061	acctccacag	tggaccttgg	gacctcaggg	actccatcct	ccctccccag	ccctacancn
	48121	netgnccctc	tctgntncc	nttaccntc	aacttnacca	taccaaacct	gcantangng
	48181	gannacatgc	nncncccnng	ntccaggaag	ttcaacacca	cngagnngt	netgcagggt
30	48241	ctgctnnnnc	ccntnttcaa	gaacnccagt	gtnggcentc	tgtactctgg	ctgcagactg
	48301	acctnnctca	ggncngagaa	gnatggngca	gccactggan	tggtatgccat	ctgcancac
	48361	cnnctnanc	ccaaaagnc	tgactgnac	agnagcngc	tnactggga	gctnagccan
	48421	ctgaccaann	ncatcnnga	gctgggnccc	tacacctgg	acaggnacag	tctctatgtc
	48481	aatgggttca	ccattggat	ccctgtgccc	accagcagca	ctcctgggac	ctccacagtg
35	48541	gaccttgggt	cagggactcc	atcctccctc	cccagcccca	caactgctgg	ccctctcctg
	48601	gtgccgttca	ccctcaactt	caccatcacc	aaactgaagt	acgaggagga	catgcatgac
	48661	cctgggtcca	ggaagtcca	caccacagag	agagtctgc	agagtctgct	tttcacctatg
	48721	ttcaagaaca	ccagtgttgg	ccctctgtac	tctgggtgca	gactgaacctc	gctcaggtcc
	48781	gagaaggatg	gagcagccac	tggagtggat	gccatctgca	cccacctgtg	tgaccccaaa
40	48841	agccctggag	tggacaggga	gcagctatac	tgggagctga	gccagctgac	caatggcatc
	48901	aaagagctgg	gtccctacac	cctggacaga	aacagtctct	atgtcaatgg	tttcacctat
	48961	cagacctctg	cgcccaacac	cagcactcct	gggacctcca	cagtgnacnt	nggnacctcn
	49021	gggactccat	cctccntccc	cngccncaca	tctgctggcc	ctctcctggt	gccattcacc
	49081	ctcaacttca	ccatcaccaa	cctgcagtac	gaggaggaca	tgcatcacc	aggctccagg
45	49141	aagttcaaca	ccacggagcg	ggtcctgcag	ggtctgctg	gtcccatggt	caagaacacc
	49201	agtgtcggtc	ttctgtactc	tggtgtcaga	ctgaccttgc	tcaggcctga	gaagaatggg
	49261	gcaaccactg	gaatggatgc	catctgcacc	cacctgtctg	accccaaaag	ccctggactg
	49321	nacagngagc	ngctntactg	ggagctnagc	canctgacca	annncatcnn	ngagctgggn
	49381	ccctacaccc	tggacaggna	cagtctctat	gtcaatggtt	tcaccatcnn	ganctctgng
50	49441	cccaccacca	gcactcctgg	gacctccaca	gtgnacntng	gnacctcngg	gactccatcc
	49501	tccntcccn	gccncacanc	nnetgnccct	ctcctgntnc	cnttcacnt	caacttnacc
	49561	atcaccaacc	tgcantangn	ggannacatg	cnncccnng	gntccaggaa	gttcaacacc
	49621	acngagaggg	ttctgcaggg	tctgctcaaa	cccttggtca	ggaatagcag	tctggaatac
	49681	ctctattcag	gctgcagact	agcctcactc	aggccagaga	aggatagctc	agccatggca
55	49741	gtggatgcca	tctgcacaca	tgcacctgac	cctgaagacc	tcggacttga	cagagagcga
	49801	ctgtactggg	agctgagcaa	tctgacaaat	ggcatccagg	agctggggcc	ctacaccttg
	49861	gaccgggaaca	gtctctatgt	caatgggtttc	acccatcgaa	gctctatgcc	caccaccagc
	49921	actcctggga	cctccacagt	ggatgtggga	acctcagggg	ctccatcctc	cagccccagc
	49981	cccacgactg	ctggccctct	cctgatacca	ttcaccctca	acttccacct	caccaacctg
60	50041	cagtatgggg	aggacatggg	tcacctgggc	tccaggaagt	tcaacaccac	agagagggtc
	50101	ctgcagggtc	tgcttgggtc	catattcaag	aacaccagt	ttggccctct	gtactctggc

	50161	tgcagactga	cctctctcag	gtctgagaag	gatggagcag	ccactggagt	ggatgccatc
	50221	tgcatccatc	atcttgacct	caaaagccct	ggactcaaca	gagagcggct	gtactgggag
	50281	ctgagccaac	tgaccaatgg	catcaaagag	ctgggcccct	acaccctgga	caggaacagt
	50341	ctctatgtca	atggtttcac	ccatcggacc	tctgtgccca	ccaccagcac	tcctgggacc
5	50401	tccacagtgg	accttggaac	ctcagggact	ccattctccc	tcccaagccc	cgcaactgct
	50461	ggccctctcc	tggtgctgtt	cacctcaac	ttaccatca	ccaacctgaa	gtatgaggag
	50521	gacatgcac	gccctggctc	caggaagttc	aacaccactg	agagggctct	gcagactctg
	50581	cttggtccta	tggtcaagaa	caccagtgtt	ggccttctgt	actctggctg	cagactgacc
	50641	ttgctcaggt	ccgagaagga	tggagcagcc	actggagtgg	atgccatctg	cacccaccgt
10	50701	cttgacccca	aaagccctgg	actgnacagn	gagcngctnt	actgggagct	nagccanctg
	50761	accaannnca	tcnnngagct	gggnccttac	accctggaca	ggnacagtct	ctatgtcaat
	50821	ggtttcaccc	atcnganctc	tgngcccacc	accagcactc	ctgggacctc	cacagtgnac
	50881	ntnggnacct	cngggactcc	atcctcctnc	cccngccnca	cancnnctgn	ccctctcctg
	50941	ntnccnttca	ccntcaactt	naccatcacc	aacctgcant	angngganna	catgcnncnc
15	51001	ccnggntcca	ggaagtcca	caccacngag	agagtctctc	agggctctgt	caggcctgtg
	51061	ttcaagaaca	ccagtgttgg	ccctctgtac	tctggctgca	gactgacctt	gctcaggccc
	51121	aagaaggatg	gggcagccac	caaagtggat	gccatctgca	cctaccgccc	tgatcccaaa
	51181	agccctggac	tggacagaga	gcagctatac	tgggagctga	gccagctaac	ccacagcatc
	51241	actgagctgg	gcccctacac	ccaggacagg	gacagtctct	atgtcaatgg	cttcacccat
20	51301	cggagctctg	tgccaaccac	cagtattcct	gggacctctg	cagtgcacct	ggaaaccact
	51361	gggactccat	cctccttccc	cggccacaca	gagcctggcc	ctctcctgat	accattcact
	51421	ttcaacttta	ccatcaccaa	cctgcgttat	gaggaataca	tgcaacaccc	tggttccagg
	51481	aagttcaaca	ccacggagag	ggttctgcag	ggtctgtcca	cgcccttgtt	caagaacacc
	51541	agtgttgccc	ctctgtactc	tggtctgcaga	ctgaccttgc	tcagacctga	gaagcaggag
25	51601	gcagccactg	gagtggacac	catctgtacc	caccgcgttg	atcccatcgg	acctggactg
	51661	gacagagagc	ggctatactg	ggagctgagc	cagctgacca	acagcatcac	agagctggga
	51721	ccctacaccc	tggataggga	cagtctctat	gtcgatggct	tcaacccttg	gagctctgtg
	51781	ccaaccacca	gcactcctgg	gacctccaca	gtgcacctgg	caacctctgg	cactccatcc
	51841	cccctgcctg	gccacacagc	ccctgtccct	ctcttgatac	cattcacctg	caactttacc
30	51901	atcacccgac	tgcattatga	agaaaacatg	caacaccctg	gttccaggaa	gttcaacacc
	51961	acggagaggg	ttctgcaggg	tctgtcaag	cccttgttca	agagcaccag	cgttggccct
	52021	ctgtactctg	gctgcagact	gaccttgctc	agacctgaga	aacatggggc	agccactgga
	52081	gtggagccca	tctgcacctc	cgccttgat	cccactgggtc	ctggactgga	cagagagcgg
	52141	ctatactggg	agctgagcca	gctgaccaac	agcatcacag	agctgggacc	ctacaccctg
35	52201	gatagggaca	gtctctatgt	caatggcttc	aaccttgga	gctctgtgcc	aaccaccagc
	52261	actcctggga	cctccacagt	gcacctggca	acctctggga	ctccatcttc	cctgcctggc
	52321	cacacaactg	ctggccctct	cctgggtgcg	ttcaccctca	acttcacat	caccaacctg
	52381	aagtacgagg	aggacatgca	ttgccctggc	tccaggaagt	tcaacaccac	agagagagtc
	52441	ctgcagagtc	tgcattggtc	catgttcaag	aacaccagtg	ttggccctct	gtactctggc
40	52501	tgcagactga	ccttgctcag	gtccgagaag	gatggagcag	ccactggagt	ggatgccatc
	52561	tgcaccaccc	gtcttgacct	caaaagccct	ggactgnaca	gngagcngct	ntactgggag
	52621	ctnagccanc	tgaccaannn	catcnnngag	ctgggncctc	acacctggga	caggnacagt
	52681	ctctatgtca	atggtttcac	ccatcnganc	tctgngccca	ccaccagcac	tcctgggacc
	52741	tccacagtgn	acntnggnac	ctcngggact	ccatcctccn	tccccngccn	cacancnnct
45	52801	gnccctctcc	tgntnccntt	cacntcaac	tnnaccatca	ccaacctgca	ntangnggan
	52861	nacatgcnnn	ncccnngntc	caggaagtcc	aacaccacng	agnngntnct	gcagggtctg
	52921	ctnnnncccn	tnttcaagaa	cncctagtgn	ggcctctgt	actctggctg	cagactgacc
	52981	tnnctcaggn	cngagaagna	tggngcagcc	actggantgg	atgccatctg	cancaccenn
	53041	cntnanccca	aaagnccctg	actgnacagn	gagcngctnt	actgggagct	nagccanctg
50	53101	accaacagca	tcacagagct	gggaccttac	accctggata	gggacagtct	ctatgtcaat
	53161	ggtttcaccc	atcgaagctc	tatgcccacc	accagtattc	ctgggacctc	tgcagtgcac
	53221	ctggaaacct	ctgggactcc	agcctccctc	cctggccaca	cagccctctg	ccctctcctg
	53281	gtgcaattca	ccctcaactt	cactatcacc	aacctgcagt	atgaggagga	catgctgcac
	53341	cctggttcca	ggaagtcca	caccacggag	agagtctctg	agggctctgt	caagcccttg
55	53401	ttcaagagca	ccagtgttgg	ccctctgtac	tctggctgca	gactgacctt	gctcaggcct
	53461	gaaaaacgtg	gggcagccac	cggcgtggac	accatctgca	ctcacccctc	tgacctctca
	53521	aaccttgagc	tgnaacngna	gcnctntac	tgggagctna	gccanctgac	caannncatc
	53581	nnngagctgg	gnccctacac	cctggacagg	nacagtctct	atgtcaatgg	tttcacccat
	53641	cnganctctg	ngccccaccac	cagcactcct	gggacctcca	cagtgnacnt	nggnacctcn
60	53701	gggactccat	cctcctctcc	cngccncaca	cnncctgncc	ctctcctgnt	ncntttcacc
	53761	ntcaacttna	ccatcaccaa	cctgcanttan	gnggannaca	tgcnnccccc	ngntccagg



	53821	aagttcaaca	ccacngagng	ngtntctgcag	ggtctgetnm	nncccntntt	caagaacncc
	53881	agtgtnggcc	ntctgtactc	tggctgcaga	ctgacctnnc	tcaggncnga	gaagnatgg
	53941	gcagccactg	gantggatgc	catctgcanc	caccnncntn	ancccaaaag	ncctggactg
	54001	nacagngagc	ngctntactg	ggagctnagc	canctgacca	annncatcnn	ngagctgggn
5	54061	ccctacaccc	tggacaggna	cagtctctat	gtcaatgggt	ttcaccctcg	gagctctgtg
	54121	ccaaccacca	gcactcctgg	gacctccaca	gtgcacctgg	caacctctgg	gactccatcc
	54181	tccttgccctg	gccacacagc	ccctgtccct	ctcttgatac	cattcacctc	caactttacc
	54241	atcaccaacc	tgcatatga	agaaaacatg	caacaccctg	gttcaggaa	gttcaacacc
	54301	acggagcggg	tcctgcaggg	tctgcttggt	cccattgttca	agaacacaa	gttcggcctt
10	54361	ctgtactctg	gctgcagact	gaccttgctc	aggcctgaga	agaatggggc	agccactgga
	54421	atggatgcca	tctgcagcca	ccgtcttgac	cccaaaagcc	ctggactgna	cagnagcng
	54481	ctntactggg	agctnagcca	nctgaccaan	nncatcnng	agctgggncc	ctacaccctg
	54541	gacaggnaca	gtctctatgt	caatgggttc	acccatcnga	nctctgngcc	caccaccagc
	54601	actcctggga	cctccacagt	gnacntnggn	acctcnggga	ctccatcctc	ontccccngc
15	54661	cncacancnn	ctgnccctct	ctgntnccn	ttcacntca	acttnaccat	caccaacctg
	54721	cantangngg	annacatgcn	ncncccnngn	tccaggaaagt	tcaacaccac	ngagnngtn
	54781	ctgcagggtc	tgctnnnncc	cntnttcaag	aacnccagt	tnggcntct	gtactctggc
	54841	tgcaactga	cctnnctcag	gncngagaag	natggngcag	ccactggant	ggatgccatc
	54901	tgcanccacc	nnentnancc	caaaagncct	ggactgnaca	gngagcngct	ntactgggag
20	54961	ctnagccanc	tgaccaannn	catcnngag	ctgggnccct	acaccctgga	caggnacagt
	55021	ctctatgtca	atggtttcac	ccatcagaac	tctgtgcccc	ccaccagtac	tcctgggacc
	55081	tcacacagt	actgggcaac	cactgggact	ccatcctcct	tccccggcca	cacagagcct
	55141	ggccctctcc	tgataccatt	cactttcaac	tttaccatca	ccaacctgca	ttatgaggaa
	55201	aacatgcaac	accctgggtc	caggaaagtcc	aacaccacgg	agagggttct	gcagggtctg
25	55261	ctcacgccct	tgttcaagaa	caccagtgtt	ggccctctgt	actctggctg	cagactgacc
	55321	ttgctcagac	ctgagaagca	ggagggcagcc	actggagtgg	acaccatctg	taccaccggc
	55381	ggtgatccca	toggacctgg	actgnacagn	gagcngctnt	actgggagct	nagccanctg
	55441	accaannnca	tcnnngagct	gggnccctac	accctggaca	ggnacagctc	ctatgtcaat
	55501	ggttttaccc	atcngancct	tgngcccacc	accagcactc	ctgggacctc	cacagtgnac
30	55561	ntnggnacct	cngggactcc	atcctccntc	cccngccnca	cancnncctg	ccctctcctg
	55621	ntnccnttca	ccntcaactt	naccatcacc	aacctgcant	angngganna	catgcnncnc
	55681	ccnggntcca	ggaagttcaa	caccacmgag	ngngtntctg	agggtctgct	nnnncccntn
	55741	ttcaagaacn	ccagtgtngg	ccntctgtac	tctggctgca	gactgacctn	nctcaggncn
	55801	gagaagnatg	gngcagccac	tggantggat	gccatctgca	nccaccnncn	tnancccaaa
35	55861	agnctggac	tgnaacngna	gncgctntac	tgggagctna	gccanctgac	caannncatc
	55921	nnngagctgg	gnccctacac	cctggacagg	nacagtctct	atgtcaatgg	tttaccatct
	55981	cggagctctg	tgccaaccac	cagcagtcct	gggacctcca	cagtgcacct	ggcaacctct
	56041	gggactccat	cctccctgcc	tggccacaca	gcccctgtcc	ctctcttgat	accattcacc
	56101	ctcaacttta	ccatcaccaa	cctgcattat	gaagaaaaca	tgcaacaccc	tggttccagg
40	56161	aagttcaaca	ccacggagag	ggttctgcag	ggtctgetca	agcccttggt	caagagcacc
	56221	agtgttgcc	ctctgtactc	tggctgcaga	ctgaccttgc	tcagacctga	gaaacatggg
	56281	gcagccactg	gagtggacgc	catctgcacc	ctccgccttg	atccacttgg	tcctggactg
	56341	nacagngagc	ngctntactg	ggagctnagc	canctgacca	annncatcnn	ngagctgggn
	56401	ccctacaccc	tggacaggna	cagtctctat	gtcaatgggt	tcacccatcn	ganctctgng
45	56461	cccaccacca	gcactcctgg	gacctccaca	gtgnacntng	gnacctcngg	gactccatcc
	56521	tccttccccn	gccncaacnc	nnctgnccct	ctcctgntnc	cnttcacctn	caacttnacc
	56581	atcaccaacc	tgcantangn	ggannacatg	cnncccnccg	gntccaggaa	gttcaacacc
	56641	acngagngng	tnctgcaggg	tctgctnnnn	cccntnttca	agaacnccag	tgtnggccnt
	56701	ctgtactctg	gctgcagact	gacctnnctc	aggncngaga	agnatggngc	agccactgga
50	56761	ntggatgcca	tctgcancca	ccnncntnan	cccaaaagnc	ctggactgna	cagnagcng
	56821	ctntactggg	agctnagcca	nctgaccaan	nncatcnng	agctgggncc	ctacaccctg
	56881	gacaggnaca	gtctctatgt	caatgggttc	acccatcgga	cctctgtgcc	caccaccagc
	56941	actcctggga	cctccacagt	gcacctggca	acctctggga	ctccatcctc	cctgcctggc
	57001	cacacagccc	ctgtccctct	cttgatacca	ttcacctcca	actttaccat	caccaacctg
55	57061	cagtatgagg	aggacatgca	tcgcccctga	tctaggaagt	tcaacaccac	agagagggtc
	57121	ctgcagggtc	tgcttagtcc	cattttcaag	aactccagt	ttggccctct	gtactctggc
	57181	tgcaactga	cctctctcag	gcccagagaag	gatggggcag	caactggaat	ggatgctgtc
	57241	tgccctctacc	accctaacc	caaaagacct	gggctggaca	gagagcagct	gtactgcag
	57301	ctaagccagc	tgaccacaaa	catcactgag	ctgggcccct	acagctctga	cagggacagt
60	57361	ctctatgtca	atggtttcac	ccatcagaac	tctgtgcccc	ccaccagtac	tcctgggacc
	57421	tcacacagt	actgggcaac	cactgggact	ccatcctcct	tccccggcca	cacancnct

57481 gncctctccc tgnntccntt caccntcaac ttnaccatca ccaacctgca ntangnggan  
 57541 nacatgcnncc ncccnngntc caggaaagttc aacaccacng agngngntct gcagggtctg  
 57601 ctnnnncccn tnttcaagaa cncagtgtn ggcntctgt actctggctg cagactgacc  
 57661 tnnctcaggn cngagaagna tggngcagcc actggantgg atgccatctg canccaccnn  
 57721 cntnanccca aaagnccctgg actgnacagn gagcngctnt actgggagct nagccanctg  
 57781 accaanmnca tcnnngagct gggnccttac accctggaca ggnacagtct ctatgtcaat  
 57841 ggtttcaccc attggagctc tgggctcacc accagcactc ctggacttc cacagttgac  
 57901 ctggaaacct cagggactcc atccccctgc cccagcccca caactgctgg ccctctcctg  
 57961 gtgcccattca ccctaaactt caccatcacc aacctgcagt atgaggagga catgcacgc  
 58021 cctggatcta ggaagttaa cgcacacag agggctcctgc agggctctgt tagtcccata  
 58081 ttcaagaaca ccagtgttg ccctctgtac tctggctgca gactgacct gctcagacct  
 58141 gagaagcagg aggcagccac tggagtggac accatctgta cccaccctgt tgatcccatc  
 58201 ggacctggac tgnacagnga gcngctntac tgggagctna gccanctgac caannncatc  
 58261 nnnagctgg gnccttacac cctggacagg nacagtctct atgtcaatgg ttccaccat  
 58321 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn  
 58381 gggactccat cctccntccc cngccnaca ncnctgncc ctctcctgnt nccnttcacc  
 58441 ntcaacttna ccatcaccaa cctgcantan gnggannaca tgcnnccccc nggntccagg  
 58501 aagttaacaa ccacngagng ngtntgcag ggtctgctnn nccccntntt caagaacncc  
 58561 agtgtnggcc ntctgtactc tggctgcaga ctgacctnnc tcaggncnga gaagnatggn  
 58621 gcagccactg gantggatgc catctgcanc caccnncntn ancccaaaag nccctggact  
 58681 nacagnagc ngctntactg ggagctnagc canctgacca annncatcnn ngagctgggn  
 58741 ccctacaccc tggacaggna cagtctctat gtcaatggtt tcacccatcg gagcttggg  
 58801 ctcaccacca gcaactcctg gacttcacca gttgaccttg gaacctcagg gactccatcc  
 58861 cccgtcccca gcccacaac tgcctggcct ctctcgtgac cattcacct aaacttcacc  
 58921 atcaccaacc tgcagtatga ggaggacatg catcgccctg gctccaggaa gttcaacacc  
 58981 acggagaggg tccctcaggg tctgcttacg ccctgttca ggaacaccag tgtcagctct  
 59041 ctgtactctg gttgcagact gaccttgctc aggcctgaga aggatggggc agccaccaga  
 59101 gtggatgctg tctgcacca nctgcctgac ccaaaaagc ctggactgna cagnagcng  
 59161 ctntactggg agctnagcca nctgaccaan nncatcnng agctgggnc ctacacctg  
 59221 gacaggnaca gtctctatgt caatggtttc acccatcnga nctctnggcc caccaccagc  
 59281 actcctggga cctccacagt gnacntnggn acctcnggga ctccatctc cntcccngc  
 59341 cncacancnn ctgnccctct cctgntnccn ttcacctca acttnaccat caccaacctg  
 59401 cantangngg annacatgcn ncnccnggn tcagggaagt tcaacaccac ngagngngtn  
 59461 ctgcagggtc tgctnnnncc cntnttcaag aacnccagt tnggcmctct gtactctggc  
 59521 tgcagactga cctnnctcag gncngagaag natggngcag ccaactggant ggaatgccatc  
 59581 tgcancacc ncnntnanc caaaagncct ggactgnaca gngagcngct ntactgggag  
 59641 ctgagccanc tgaccaann catcnngag ctgggncct acacctgga caggnacagt  
 59701 ctctatgtca atggtttcc ccattggatc cctgtgccc ccagcagcac tctgggacc  
 59761 tccacagtgg accttgggtc agggactcca tctcctccc ccagccccc aactgctggc  
 59821 cctctcctgg taccattcac cctcaacttc accatcacca acctgcagta tggggaggac  
 59881 atgggtcacc ctggctccag gaagtcaac accacagaga gggtcctgca ggtctctgt  
 59941 ggtcccatat tcaagaacac cagtgttggc cctctgtact ctggctgag actgacctct  
 60001 ctcaggtccg agaaggatgg agcagccact ggagtggatg ccatctgcat ccatcatctt  
 60061 gaccccaaaa gccctggact gnacagnag cngctntact gggagctnag ccanttgacc  
 60121 aannncatcn nngagctggg nccctacacc ctggacaggn acagtctcta tgtcaatggg  
 60181 ttcacccatc nganctctgn gccaccacc agcactcctg ggacctccac agtgnacntn  
 60241 ggnacctcng ggactccatc ctcntcccc ngccnccan cnnctgnccc tctcctgntn  
 60301 ccnttcaccn tcaacttnac catcaccaac ctgcantang nggannacat gcnnccccc  
 60361 ggnccagga agttcaacac cacngagngn gtntctcagg gtctgctnnn nccntnttc  
 60421 aagaacncca gtgtnggcn tctgtactct ggctgcagac tgacctnnt caggncngag  
 60481 aagnatggng cagccactgg antggatgcc atctgcancc accnncntna ncccaaaagn  
 60541 cctggactgn acagnagcn gctntactgg gagctnagcc anctgaccaa nnnatcnnn  
 60601 gagctgggnc cctacacct ggacaggnac agtctctatg tcaatggttt caccatcag  
 60661 accttgcgc ccaacaccag cactcctggg acctccacag tggaccttgg gacctcaggg  
 60721 actccatcct cctcccccag cctacatct gctggcctc tctgggtgac tctcaccctc  
 60781 aacttcacca tcaaccaact gcagtacgag gaggacatgc atcacccagg ctccagggaag  
 60841 ttcaacacca cggagcgggt cctgcagggt ctgcttgggt ccatgttcaa gaacaccagt  
 60901 gtcggccttc tgtactctgg ctgcagactg acctgtctca ggcttgagaa gaatggggca  
 60961 gccaccagag tggatgctgt ctgacccat cgtcctgacc ccaaaagccc tggactgnac  
 61021 agngagcngc tntactggga gctnagccan ctgaccaann ncatcnnga gctgggnc  
 61081 tacacctgg acaggnacag tctctatgtc aatggtttca cccatcnan ctctgngccc

	61141	accaccagca	ctcctgggac	ctccacagtg	nacntnggna	cctcngggac	tccatcctcc
	61201	ntccccngcc	ncacagcccc	tgtccctctc	ttgataccat	tcaccctcaa	ctttaccatc
	61261	accaacctgc	attatgaaga	aaacatgcaa	cacctgggtt	ccaggaagtt	caacaccacg
	61321	gagagggttc	tgcagggtct	gtcraagccc	ttgttcaaga	gcaccagcgt	tggccctctg
5	61381	tactctggct	gcagactgac	cttgctcaga	cctgagaaac	atggggcagc	cactggagtg
	61441	gacgccatct	gcaccctccg	ccttgatccc	actggtcctg	gactggacag	agagcggtta
	61501	tactgggagc	tgagccagct	gaccaacagc	gttacagagc	tgggccccta	cacctgggac
	61561	agggacagtc	tctatgtcaa	tggtctcacc	cagcggagct	ctgtgccaac	caccagtatt
	61621	cctgggacct	ctgcagtgca	cctggaaaac	tctgggactc	cagcctccct	ccctggccac
10	61681	acagccccctg	gccctctcct	ggtgccattc	accctcaact	tcactatcac	caacctgcag
	61741	tatgaggtgg	acatgcgtca	ccctgggttc	aggaagttca	acaccacgga	gagagtctctg
	61801	caggggtctgc	tcaagccctt	gttcaagagc	accagtgttg	gccctctgta	ctctggctgc
	61861	agactgacct	tgctcaggcc	tgaanaacgt	ggggcagcca	ccggcgtgga	caccatctgc
	61921	actcaccgcc	ttgacctctc	aaaccttgga	ctggacagag	agcagctata	ctgggagctg
15	61981	agcaaaactga	cccgtggcat	catcgagctg	ggcccctacc	tccctggacag	aggcagttctc
	62041	tatgtcaatg	gtttcaccct	tcggaaactt	gtgcccatca	ccagcaactc	tgggacctcc
	62101	acagtaacac	taggaacctc	tgaactccca	tcctccctac	ctagaccctc	agtgccctggc
	62161	cctctccctgg	tgccattcac	cctcaacttc	accatcacca	acttgcatga	tgaggaggcc
	62221	atgcgacacc	ctggctccag	gaagttcaat	accacggaga	gggtccctaca	gggtctgtctc
20	62281	aggcccttgt	tcaagaatac	cagtatcggc	cctctgtact	ccagctgcag	actgaccttg
	62341	ctcaggccag	agaaggacaa	ggcagccacc	agagtggatg	ccatctgtac	ccaccacctc
	62401	gacctcctaaa	gccctggact	gaacagagag	cagctgtact	gggagctgag	ccagctgacc
	62461	cacggcatca	ctgagctggg	ccctacacac	ctggacaggg	acagtctcta	tgtcgatggt
	62521	ttcactcatt	ggagccccat	accgaccacc	agcactcctg	ggacctccat	agtgaacctg
25	62581	ggaacctctg	ggatcccacc	ttccctccct	gaaactacan	cnctgncctc	tctcctgntn
	62641	ccnttcacn	tcaacttnac	catcaccaac	ctgcantang	nggannacat	gcnnccccn
	62701	ggntccagga	agttcaacac	cacngagagg	gttctgcagg	gtctgtctcaa	gcccttgttc
	62761	aagagcacca	gtgttgcccc	tctgtattct	ggctgcagac	tgacctgtgt	caggcctgag
30	62821	aaggacggag	tagccaccag	agtggacgcc	atctgcaccc	accgcccctga	ccccaaaatc
	62881	cctgggctag	acagacagca	gctatactgg	gagctgagcc	agctgaccca	cagcatcact
	62941	gagctgggac	cctacaccct	ggatagggac	agtctctatg	tcaatggttt	caccagcggg
	63001	agctctgtgc	ccaccaccag	cactcctggg	actttcacag	tacagccgga	aacctctgag
	63061	actccatcat	ccctccctgg	ccccacagcc	actggccctg	tctctgtgcc	atccacctc
35	63121	aatttttacca	tcactaacct	gcagtatgag	gaggacatgc	atcgccctgg	ctccagggaag
	63181	ttcaacacca	cggagagggg	ccttcagggt	ctgcttatgc	ccttgttcaa	gaacaccagt
	63241	gtcagctctc	tgtactctgg	ttgcagactg	accttgcctc	ggcctgagaa	ggatggggca
	63301	gccaccagag	tggatgtgtg	ctgcacccat	cgctctgacc	ccaaaagccc	tggactggac
	63361	agagagcggc	tgtactggaa	gtcgagccag	ctgacccacg	gcactcactga	gctgggcccc
40	63421	tacaccctgg	acaggcacag	tctctatgtc	aatgggttca	cccatcagag	ctctatgacg
	63481	accaccagaa	ctcctgatac	ctccacaatg	cacctggcaa	cctcgagaa	tccagcctcc
	63541	ctgtctggac	ctacgaccgc	cagccctctc	ctggtgctat	tcacaattaa	cttcaccatc
	63601	actaacctgc	ggtatgagga	gaacatgcat	caccctggct	ctagaaagtt	taacaccacg
	63661	gagagagtcc	ttcagggtct	gctcaggcct	gtgttcaaga	acaccagtgt	tggccctctg
	63721	tactctggct	gcagactgac	cttgctcagg	cccaagaagg	atggggcagc	caccaaagtg
45	63781	gatgccatct	gcacctaccg	ccctgatccc	aaaagccctg	gactggacag	agagcagcta
	63841	tactgggagc	tgagccagct	aaccacagc	atcactgagc	tgggccccta	cacctgggac
	63901	agggacagtc	tctatgtcaa	tggtttcaca	cagcggagct	ctgtgccacc	cactagcatt
	63961	cctgggaccc	ccacagtgga	cctgggaaca	tctgggactc	cagtttctaa	acctggtccc
50	64021	tcggctgcca	gccctctcct	ggtgctatcc	actctcaact	tcaccatcac	caacctgcgg
	64081	tatgaggaga	acatgcagca	ccctggctcc	aggaagttca	acaccacgga	gagggtccct
	64141	cagggcctgc	tcaggtccct	gttcaagagc	accagtgttg	gccctctgta	ctctggctgc
	64201	agactgactt	tgctcaggcc	tgaanaaggt	gggacagcca	ctggagtggg	tgccatctgc
	64261	accaccacc	ctgaccccaa	aagccctagg	ctggacagag	agcagctgta	ttgggagctg
	64321	agccagctga	cccacaatat	cactgagctg	ggccactatg	ccctggacaa	cgacagcctc
55	64381	tttgtcaatg	gtttcactca	ctggagctct	gtgtccacca	ccagcactcc	tgggaccccc
	64441	acagtgtatc	tgggagcctc	taagactcca	gcctcgatat	ttggcccttc	agctgccagc
	64501	catctcctga	tactattcac	cctcaacttc	accatcacta	acctgcggta	tgaggagaac
	64561	atgtggcctg	gctccaggaa	gttcaacact	acagagaggg	tccttcaggg	cctgctaagg
60	64621	cccttgttca	agaacaccag	tggtggccct	ctgtactctg	gctccaggtc	gaccttgtctc
	64681	aggccagaga	aagatgggga	agccaccgga	gtggatgcca	tctgcaccca	ccgcccctgac
	64741	cccacaggcc	ctgggctgga	cagagagcag	ctgtatttgg	agctgagcca	gctgaccac

64801 agcatcactg agctggggccc ctacacactg gacaggggaca gtctctatgt caatgggtttc  
 64861 acccatcgga gctctgtacc caccaccagc accgggggtg tcagcgagga gccattcaca  
 64921 ctgaacttca ccatcaacaa cctgcgtac atggcgagca tgggccaacc cggctccctc  
 64981 aagttcaaca tcacagacaa cgtcatgaag cacctgctca gtcccttgtt ccagaggagc  
 5 65041 agcctgggtg caggttacac aggtgcagg gtcacgcac taaggctctgt gaagaacggt  
 65101 gctgagacac ggggtggacct cctctgcacc tacctgcage ccctcagcgg cccaggtctg  
 65161 cctatcaagc aggtgttcca tgagctgagc cagcagaccc atggcatcac cggctgggc  
 65221 ccctactctc tggacaaaga cagcctctac cttaacggtt acaatgaacc tggctagat  
 65281 gagcctccta caactcccaa gccagccacc acattcctgc ctctctctgc agaagccaca  
 10 65341 acagccatgg ggtaccacct gaagaccctc aactcaact tcaccatctc caatctccag  
 65401 tattcaccag atatgggcaa gggctcagct acattcaact ccaccgaggg ggtccttcag  
 65461 cacctgctca gaccttgtt ccagaagagc agcatgggccc ccttctactt ggggtgccaa  
 65521 ctgatctccc tcaggcctga gaaggatggg gcagccactg gtgtggacac cacctgcacc  
 65581 taccaccctg accctgtggg cccgggctg gacatacagc agctttactg ggagctgagt  
 15 65641 cagctgaccc atggtgtcac ccaactgggc ttctatgtcc tggacaggga tagcctcttc  
 65701 atcaatggct atgcacccca gaatttatca atccggggcg agtaccagat aaattccac  
 65761 attgtcaact ggaacctcag taatccagac ccacatcct cagagtacat caccctgctg  
 65821 agggacatcc aggacaaggt caccacactc tacaaggga gtcaactaca tgacacattc  
 65881 cgcttctgcc tggtcaccaa ctgacgatg gactcogtgt tggctactgt caaggcattg  
 20 65941 ttctcctcca atttggacc cagcctgggt gagcaagtct ttctagataa gacctgaa  
 66001 gcctcattcc attggtggg ctccacctac cagttgggtg acatccatgt gacagaaatg  
 66061 gagtcacag tttatcaacc aacaagcagc tccagcacc agcacttcta cctgaatttc  
 66121 accatcacca acctaccata ttcccaggac aaagcccagc caggcaccac caattaccag  
 66181 aggaacaaaa ggaatattga ggatgcgctc aaccaactct tccgaaacag cagcatcaag  
 25 66241 agttattttt ctgactgtca agtttcaaca ttcaggtctg tcccaacag gcaccacacc  
 66301 ggggtggact cctgtgttaa ctctcgcca ctggctcgga gagtagacag agttgccatc  
 66361 tatgaggaat ttctgoggat gaccgggaat ggtaccagc tgcagaactt caccctggac  
 66421 aggagcagtg tcttgtgga tgggtattct ccaacagaa atgagccctt aactgggaat  
 66481 tctgaccttc cctctgggc tgtcactctc atcggtctgg caggactcct gggactcatc  
 30 66541 acatgcctga tctgcggtgt cctggtgacc acccgccggc ggaagaagga aggagaatac  
 66601 aacgtccagc aacagtgcc aggtactac cagtcacacc tagacctgga ggatctgcaa  
 66661 tgactggaac ttgcccgtgc ctgggggtgc ttccccccag ccagggtcca aagaagcttg  
 66721 gctggggcag aaataaacca tattgttcgg aaaaaaaaaa aaaaa

## 35 SEQ ID NO. 3

hk5 amino acid

MATARPPWMWVLCALITALLLGVTEHVLANNVSCDHPSNTVPSSGNQDLGAGAGEDARSDDSSSRIINGS  
 40 CDMHTQPWQAALLRPNLQYCGAVLVHPQWLLTAHCRKKVFRVRLGHYSLSPVYESGQMQFGVKSIPHPG  
 YSHPGHSNDLMLIKLNRIRPTKDVREINVSSHCPASGATKCLVSGWGTTKSPQVHFVKVLQCLNISVLSQKR  
 CEDAYPRQIDDTMFCAGDKAGRSDSCQGDSSGPPVVCNGLSLQGLVSWGDYPCARPNNRPGVYTNLCKFTKWIQET  
 IQANS

## 45 SEQ ID NO. 4

KLK5 CDS

ggtgtctgtg cgtcctgcac ccacatcttt ctctgtcccc tccttgccct gtctggaggc  
 tgctagactc ctatcttctg aattctatag tgcctgggtc tcagcgcagt gccgatggtg  
 gcccgctcct gtggttccto tctacttggg gaaatcaggt gcagcggcca tggctacagc  
 50 aagaccccc tggatgtggg tgctctgtgc tctgatcaca gccttgcttc tgggggtcac  
 agagcatggt ctgcacaaca atgatgtttc ctgtgaccac ccctctaaca cctgcccctc  
 tgggagcaac caggacctgg gagctggggc cggggaagac gcccggtcgg atgacagcag  
 cagccgcac atcaatggat ccgactgcga tatgcacacc cagccgtggc aggcgcgct  
 gttgctaagg ccaaccagc tctactgcgg ggcggtgttg gtgcatccac agtggctgct  
 55 cagggccgcc cactgcagga agaaagtttt cagagtccgt ctgggccact actccctgtc  
 accagtttat gaatctgggc agcagatgtt ccaggggtc aaatccatcc cccaccctgg  
 ctactccac cctggccact ctaacgacct catgtcatc aaactgaaca gaagaattcg  
 tccactaaa gatgtcagac ccatcaacgt ctctctcat tgcctctctg ctgggacaaa

```

gtgcttggtg tctggctggg ggacaaccaa gagcccccaa gtgcacttcc ctaaggtcct
ccagtgcctt aatatcagcg tgctaagtca gaaaagggtc gaggatgctt acccgagaca
gatagatgac accatgttct gcgcccgtga caaagcaggt agagactcct gccagggtga
ttctgggggg cctgtggtct gcaatggctc cctgcaggga ctctgtctct ggggagatta
5 ccttgtgccc cggcccaaca gaccgggtgt ctacacgaac ctctgcaagt tcaccaagtg
gatccaggaa accatccagg ccaactcctg agtcatccca ggactcagca caccggcatc
cccacctgct gcagggacag cctgacact cctttcagac cctcattcct tcccagagat
gttgagaatg ttcatctctc cagccctga ccccatgtct cctggactca gggctgtgctt
ccccacatt gggctgaccg tgtctctcta gttgaaccct gggaacaatt tccaaaactg
10 tccaggggcg gggttgctc tcaatctccc tggggcactt tcctcctcaa gctcaggggc
catcccttct ctgcagctct gacccaaatt tagtcccaga aataaactga gaagtggaaa
aaaaaaa

```

## 15 SEQ ID NO. 5

KLK5 nucleic acid

```

gggcccagag tgaaggcaag agaaggagtt gagagctccc tctgcaaagt ggcttgagtc
tcccctgcct aaaatgcagg gagagggagg cagaaagaca ggggaagagga aggggtgggg
20 aagaaaagaga gagagagaga gagacagaat aacacaacta cagaaacaca gagagaacac
acagagagcc tgggacacag ggacacacag agtcagagag aaaagagaag atagagaaaag
acacaaatgg agacacagag gtgtaaagaa agagagatta acagagtccc agatacacgc
aaagggcgag aagcacagtt ttcagggtgg tgtctatgat catctctctt tttttttttt
tttttttttt tttttgagac ggagtctctc tctgtcgcct aggtctggagt gcagtggcgg
25 gatctcggct cactgcaagc tccgcctccc gggttcacgc cattctctg cctcagcctc
ccaagtagct gggactacag gcgcccgcga ctacgcccgg ctaatttttt tgtattttta
gtagagacgg ggtttcaccg ttttagccgg gatggcctcg atctcctgac ctctgtatcc
gcccgcctcg gcctcccaaa gtgctgggat tacaggcctg agccaccgag cccggccatg
atcatcttct tgactatgct gatgtgacaa gtacctaaag ccatcagact ctacccttta
30 aatatgcagt ttgggccagg caccgtggct catgcctgta attccagcac tttgggaggc
agagggtgggt gaatcacttg aggcaggag tttgagacca gcctggccaa catggtgaaa
ctctgtcttt actaaaaaaa aaaaaaaaaa aaaaaaaatc agccgggtgt cgtggggcac
acctgtaatc ccagctatgc tggaggtgga ggcacgagag tcacttgaa cctggaggcg
gaggttgtag tgggccgaga tcacatcacc gccctccagc ctgggcgaca gagcaagact
35 ctgtctcaaa taaataaata aacaaacgaa caagcagttt gttgtacctt agttatatct
aaaaaaaaaa tgctgtcaac aaatagagca gaagtgaat aaaggaaaaat aaatgggcca
agaactctaa ggtatatattg acaaatcatt cagaaccttt aaaaaagaaa gaatcacaga
ggcatagaaa gacagggagg aacagggaga cagaaacacc tgtggcccaa ggagaacaaa
acaaggctcc taagacagac aggagggagag agagagagag tgagtggagag acagacagag
40 aaaaagacag agagagagag acagagacag agagacagag aggcgagagg gatagaaaaga
gagagagggg tggagagaga cagagatat tgagagagac tcagaaagat agccgagggg
gaaccacaga gagatggaag aagactctga gaaaaaacca gagacaaaga tggaaaagag
agtatcgagg gtgaacagac agtggtgga tgagcaaaat gcagagaaga aagcaagcaa
tccaggcgcc aagaatagtg acccagagtt ggtgagaagc cagatcctta aggctggggg
45 aggcagggaa ggggctggcc tggcttccgg agacctctcc ccattctccg ggccaggagg
gtaggggagtg acattccgga ctgggtgggg ggtgctctgg ggtggagat agggggagca
ggaggagcta ttgctaaggc ccgataggca cctcattgcc cgggaatgtg cccagggag
cagtgggtgg ttataactca ggcccgtgac ccagagccca ggaggaggca gtggccagga
aggcacaggc ctgagaagtc tgcggtgag ctgggagcaa atccccacc ccctacctgg
50 gggcaggggc aagtgaagac tggtaggggt ggctcagcag gcagggaagg agaggtgtct
gtgcgtcctg cacccacatc tttctctgtc ccctccttgc cctgtctgga ggctgtaga
ctcctatctt ctgaattcta tagtgcttgg gtctcagcgc agtgccgatg gtggcccgct
cttgtggttc ctctctacct ggggaaataa ggtaggggag ggaggggaag tgggttaagg
gctccccgga tcgcctgggc ctccaacccc totgacattc cccatccagg tgcagcgcc
55 atgggtacag caagaccccc ctggatgtgg gtgctctgtg ctctgatcac agccttgcct
ctgggggtca caggtaacca gaactctggg gtgggagggg tgtgggattg ggaggactgt
ctctcgggca ctagagcgcc tgtcccctgg ggaactgtgt gagcctgggc atgactcgg
gaccgggtga atgtgagtct ctgtctgtac ttgtggtgtg gcgatcgtat gtggccctgt

```

	gactgccacg	gtgtgtgtcg	gggaggggga	tgccttttcc	catatcaggt	gactgtgchg
	cagggtggcac	tgaccctttg	aggctgtgtg	tgtgggtttg	tgattgtgtg	tgcatttaag
	attgtgtgtg	gctccacagc	tgtgtgggtg	aatgcatgta	gcactggggg	tgttcactgt
	gtgtttggct	gtgtgtgggt	acttggcatt	gtatatgact	gcaggtatct	gcagttcctg
5	tccctgaggt	cccgggattg	cgtgcaacaa	aagtggcat	caccatggaa	agctgtgact
	gtgtgctgct	tgccaggcgat	tatgtgattg	tggctgagtg	tgacgttatg	gatgcccgta
	tttgtgaccg	tgtgactacc	tgaagctctg	tgtaggggtg	actgtatgtg	actgtgtgtg
	tctgtgtgag	gccgtgtaaa	tgctactgta	tgtgtgatgg	tgacgtgtg	tgctgtgagt
	ttctgtctct	gcctggaggg	atagagggtg	caggggtagc	tatctctggg	agatgggtgc
10	cagggtgactg	acttgcagtg	tgtgcctgtg	tgccagaagag	tatgtggcag	tctgaacatc
	tgtgcacaca	cggcatctgt	gcgtggcact	gagacactgt	ggatgagggg	gtgcgatccc
	gctaggctgc	ccgggagcgt	gtgtacctgg	agacagagct	gtatgttagc	tgccacctgtg
	gagccaacat	ggcgctgtct	gcagaactgc	gtgcgtgctt	ggctgttact	gctgttgtgc
	gcgtggttct	tggggtgagt	tcgtgaatga	tggtgggtcc	agggccatca	gcaagggtaa
15	gaaccaggcc	ggcgcggtg	gctcacgcct	gtaatcccag	ccctttggga	ggccgaggca
	ggcggtatcac	ctgaggtcgg	gagatcgagg	ccagcctgac	caacatggag	aaccccgctc
	ctactaaaaa	tacaaaaaat	tagctgggtg	ggtggcgctg	gcctgtaatc	ccagctactc
	gggagactgg	ggcagaaaaa	tcgcttgaac	ccgggaggtg	gaggttgchg	tgagccgaga
	tcgcgccatt	gcactccagc	ctgggcaaca	agagcgaaac	tccgtctcga	aagaaaaaaa
20	gaaaaaaaaa	agggtaaaga	ccagtgaatg	ggcacgggag	gactgatgat	ggagtggggc
	atgcatgtag	tctgtaggtc	tgtgtgtgag	aggaggagat	tgacaggatt	gagaaggcat
	gttttcatct	gagaattcag	aaacctaggg	ctgctcttcc	cctccatgtg	gccccctaag
	ctgagccctt	ctttcctggg	cctgctttcg	gaaccctagc	tccgccccatg	agctctgacc
	ccacctcctt	tcctcaacca	cgcctctagg	ccagactcta	gtggaccccg	cctaaggcca
25	caccctcttg	ggccaggctc	caccctctat	tctgtgggta	ccttctagaa	cccccttcaa
	agtccagagct	tttttttttt	tttttttgga	gacagtcttg	ctctctctcc	caggctggag
	tgccagtggcg	tgatctcgcc	tcactgcaac	ctctgcctcc	caggttcaag	tgattctcgt
	gcctccacct	cctgagtagc	tgggattaca	ggtgcgcgcc	accacgcctg	gctaattttt
	gtgtcttttag	tagagacagg	gtttcacctt	gttgccagg	ctggtotcaa	actcccaacc
30	tcagggtgatc	cgccccacctc	ggcctcccag	agtgtctggg	ttacaggcgt	gagccaccgc
	cccagaccca	aagtcagagc	tctttatagg	agactctaac	atgtaaccct	gaccctggcc
	ctaactaagt	caattccaaa	ccccctcctg	cctccagccc	tgacccccact	cactgaggcc
	tgacccccact	tcttgagacc	agttccatcc	ctaaaagccct	ggtctccctc	ccatccccag
	gtccacagccc	ccacagcttt	ggcactaccc	ctgagcttgt	ccaggaaatcc	tgtacccaat
35	tttaccctca	catgtagtcc	tagccaattc	caggaaatctg	tgagggtccag	ttagagtcca
	gtaaccctac	ctgagcctgg	gctctgtcct	tgagcttgag	cctgggcttg	agaggtgcca
	ctcttattct	ccaggccctg	ccccctgccc	ctcagcatgt	cagacaccca	cctctagct
	ggtctggcct	cttgagtctg	aaaccacccc	ccagcccaag	ccccgcctct	gagccccgcc
	caaccctatt	tcggttccca	gagcatgttc	tcgccaacaa	tgatgtttcc	tgtgaccacc
40	cctctaacac	cgtgccctct	gggagcaacc	aggacctggg	agctggggcc	ggggaagacg
	ccgggtcgga	tgacagcagc	agccgcacat	tcaatggatc	cgactgcgat	atgcacaccc
	agccgtggca	ggccgcgctg	ttgctaaggc	ccaaccagct	ctactgcggg	gcgggtgttg
	tgcatccaca	gtggctgctc	acggccgccc	actgcaggaa	gaagtgtagt	ggagtcccaa
	gaggagggtt	ggtggggacg	gggaagtggg	ggtgggggtg	gggaagtggg	ggtgggggtg
45	tcatggaggt	gagggctggg	ggggacgggg	aagtgggggt	gggggtgtca	tggaagggtga
	gggttggtgg	ggatgggttg	gggatgtggg	agcaggagga	ggtcaggttg	gggataggac
	taaggatgga	gttttgcggg	ggagcaagg	gggaggtatg	ggttgagag	gggagagtgt
	tgtggttagg	aatgggaagg	agccaaggat	gggttggtat	tggggttagg	agcatatatt
	tggtgaatgg	tttgggatgg	aggtggaatt	gggattggct	ttagaattgg	gggtgggtga
50	aaatcgggct	gggggtgaaa	tgaagatagc	atggagatag	ggttgagatt	gggagcagat
	atagaatgaa	ggatggggat	tggagtgttg	ggtgggggtg	gagatggttg	gatttgggct
	tgagaatgca	tatgtgtgat	gcttctgggt	aggaaagaa	ttagggttg	gaatgggatg
	ggtttggaat	tgtgactggg	atggggacag	gcattggggt	ggagaccaag	agggagttag
	ggatgggttg	gggaccgggg	gtggggatgg	gggtggggct	ggggctgggt	gtgggggttg
55	gattggcggt	ggacgtggag	atagagatca	gggttggttg	tgacctgccc	catctcctc
	agagttttca	gagtcctgct	cgccacttac	tcctgtctac	cagtttatga	atctggggcag
	cagatgttcc	agggggtcaa	atccatcccc	caccctggct	actccacccc	tgccactct
	aacgacctca	tgctcatcaa	actgaacaga	agaattcgct	ccactaaaga	tgccagaccc
	atcaacgtct	cctctcattg	tcctctgct	gggacaaagt	gcttggtgtc	tggtctgggg

5 acaaccaaga gcccccaagg tgagtgtcca ggttcttctt gataccgacc catctctgcc  
gcctttccatc tttctccact tctcattgtg ttctgttttg acagtgcact tccctaaggt  
cctccagtg c ttgaatatca gcgtgctaag tcagaaaagg tgcgaggatg cttacccgag  
acagatagat gacaccatgt tctgcgcggg tgacaaagca ggtagagact cctgccaggt  
gaggacacct ctcttttattc agcagataca cactgagtgc caactcggta acatggagcg  
10 ttgccaaatt ctgagaatcc agcaattgcc aagacagtca ggacccctgt tctcacagag  
ctcataccct agagtagtgg tgtttagtag aaataatgct gagctgttta tgtcatttcc  
agtttttttag tagccacatt aaaacaggta aaaaaggctg ggcgcagtgg ctcacacctg  
taatcccagc actttgggag gctgaggcag gcagatcacc tttggtcagg agtttgagac  
tagcctggcc aacatggcga aactctgtct ctaaaaaaa atacaaaaat tagcctggca  
tggtggcggg cgctgttaat ctcagctgct caggaggccg agacacaaga atcacttaaa  
cccaggaggt ggaggttgca gtgagctgag atcgtgccac tcaactccaa cttgggagaca  
gagtgacact tttgtctcaa aaagaaaaaa aaaaacaagt aaaaaagaaa caggtgaagt  
taactttaat aacccaatgt atcccaaata caatcatttc aaagtgaat taatataaaa  
15 caattatgaa tgagatactt tacattcttt tcttgttttc atattaagtc tttgaaagt  
agtatatatg ttatgctgac agcacatctc aatttggact agctacattt caggtgctca  
gtagccacat gtggctagca gttactgtat tggatggcac ggatctagag ggaagatca  
gggctgtttt gtatggttgg gcaggttgtg cactgcataa agataccata tctaataagg  
gcaactccgtg ttacagatgt cagttttggc agttttcagg cgtgtggtag ttaagtgtct  
20 tgtttcaaca aaatctgtaa tatgacagtt ttctagcaag tgctggtaaa atatcttgag  
gaaggaaaag agaaatctgg taggtatttt tacaagagaa tatttaatac aggggattaa  
ttgcaaagct gctggaaggg ctggaggaac aaagttaaaa aataaaaaac tctgtggtca  
agaatctgca taaatagggc aatttcagag agtggtaaag gtttaaccca aaataaaaa  
tggttttagg atagtaaaaca ataaggcca atattcaaaa aggtgttcag gggagcctcc  
25 ttggagaggt ggcatttgag cagagaatgg atgacacaaa gaagctaaac tctgtaagtt  
taaggggaaa gaaaaggcac gtgcaaaggc cctgaggcag taaggaaattt ggctgattca  
aagaagaaga ggaaccaaat gcaactggag acaaaaagtg ggggcaacag tagaaagtga  
cgctggaggt gtaggcaggg gcgaatgctc tgcaagtatt tcttggtcac caacacagag  
cttccctatg ttctaattga agctgtatct gttgaggaag acagaattta aatcaaaact  
30 gttacatcaa ccagcacctt tctctgtatt caggctccca agggatctag aaggacgtaa  
gttaacaagc tctcattagc aggggtgtgtg tttcaacagt agttaggaag cttgggattc  
aggagtactc cagtcoccatg gctatgaaaa gctcccccca aattgtacaa acctgacaaa  
tgcaacacct cccagctctt cccatttct tctctgtgcc ctgggtgtgg gggggtgggt  
35 tgcgaggggg aaaaactttta acagaagaaa gcacatctcg gccgggctgt gtggtcaca  
cctgtaatcc caacactttg ggaggccgag gcgggtggat cactaggtca ggagatggag  
accatcctgg ctgacacggt gaaaccctgt ctctactaaa aacacaaaaa attagccggg  
cgtggtggca ggcgcctgta gtcccagcta ctggggaggc tgaggcagga gaatggcctg  
aaccggggag gcggaacttg cagtgaagcg aggttgacc actgcactcc agcctgggca  
acacagttag actcgtctc aaaaaaaa aaagaaaaga aaagaaatca catctcattc  
40 aagtgtggc atttaaaact atttagcctt tctgtaggca aggttagtat cttgtttttc  
cagacctcaa ggtgtttttt tgtttgtttt ttcataccgg tgtgtgtct ggggtggcc  
actaaaagct acaagcaaga aataataaca actacaacaa tactaatacc aatagtataa  
aaataatagc atctggctaa ttgttgaca ctgttttaag tggtttgcat gcctcagctc  
45 attaaactcat ttacctgtta ttattggccc tattttacaa acaaggagcc aaggctcaga  
gcagttaact aacagcctct caaaagaaac tctgcagaga tattaattt aaaaaataat  
gagagaaatt aaaccacaag aaagttgaaa tttagaggta caggcagcta agcttgtttg  
ctttgaaaca gtgtctgcta ctgggaaaaa ggcaagtctt ggctttccta ataattgata  
ccaggactct gtaattcata ttttgcatgc atgtaagtaa gaaatgaagc cgggtgcaat  
50 ggcacatgcc agtaatccca gcaacttaaa attaaaaaaa taaaaatact aattgtttt  
ggagtccaag accagcctgg gcaacttaaa atcattattg tagtatgtac atatttttt  
attttagtag attttattca taccacttac acggagtctc gctctgtcac ccaggctgga  
cttttctttt cttttctttt cttttttgag catgcccctc ctgggttcaa gcattttctt  
gtgcaatggc accatatcag ctgggataac aggcaccac caccatgcct ggctattttt  
55 cacctcagcc tccaagtag ttccaccatg ttggccaggc tacaggtgtg agccaccgtg  
ttttttccgt agagatggg gcctcccaaa ttgctggtat ggtccactca agctactttt  
ccagtgatct gcctgcctcg cctcccaaa tttctcttac ctcaaacaga caggaaatgg  
cccaggtggg agatagacat gccgagtggc taattttgcac cggcgtaggac ctcaggttcc  
cattttcttc atcaatatta aacagaccaa aaacttctg cgcgtaggac ctcaggttcc

ccaagtggaa gacaggcaat aaagagatag ataaatatgt agtaaattaa aaaaaaaaaa  
 aattagccgg gtgtggtggc ttgcacctgt agttccagct acttgggagg ctgaggtggg  
 agaattgctt gagcccaaac gtttgaggct gcggtaagcc atgactgcac tgctgcactc  
 cagacagcag cctgggtgac aaagcaagac gtttttgtca gaaagaaaaa aaaaagagac  
 5 gaagggaggga aggagagaga aaggaaggaa ggaaggagaa agaaaggaa gaaaggagaa  
 gaaaggaaag aaggaaggag aaagaaaggaa agaaagagaa agaaagaaaa agaaagaaag  
 aaagaagaaa gaaaagagag aggaaggaa gaaagaaggaa aaagaggaa aaaaatgact  
 gttgaagagc agtgagtatt attataggag ggtaattata gggaggtatg ggggaattgaa  
 gacaggaaac acaaatagtt ccaagcgaat ggattttctat tgggagtgat tctgcccta  
 10 gaagacactg gcaataccag gagacatttt tggttgtcac aactatatgg aggggcatta  
 ctggcaacta atggatagat gccaaagtgt ctgttcaaca tgctatgatg cacacggcag  
 gctccacaa caaaccatta tccagcttca gatgccaca gtgccagat cgaggaaccc  
 tcatccagg gctgagaacc gtatttttgc agaaggagg tataaggatg ggttgggtgga  
 gaattgggaa ggaagggtgtg ttgccagtaa gagaaataag gcctgcacag gctggagggg  
 15 agagtggag agaaaggag gcggagagat acacgatgag ggagacaggc tggaaacagaa  
 agtagagacg aagattcgag atgtggagag gaagggtcac agacccccc gaaatgatgt  
 gtggacaaca ggaatctgga agaggaagat ggagtggaga gtgacaaatg ggttctaaaag  
 gttgaacttg gaggccaggc atgggtggct acgcctgtaa tcccaacact ttggaggctg  
 aggtggggcga atcacttgag gccaggagtt cgagaccagc ctggccaaca tggtgaaacc  
 20 cgtctctac aaaaaaata caaaaaatta gccgggtgtg gtgatggaca cctgtagtca  
 cagctacttg ggaggctgag gcaggagaat tgcttgaacc cgggagatgg aggtgcagt  
 gagctgaggt caggccactg cgctccaacc tgggcaacag agtaagactc catctcaaaa  
 aaaaaaaagc tggatttga gtgaaatatt aataacattc tccctctctc tccctttgcc  
 tgtgtctoca tctctgtctt tttctgcatt tcttcatctc tgtactttcc atctctgtg  
 25 gtctgttccc atctgtctct ccatctatgg gcatctctgg gtctctcatg tctctctctg  
 cccactttgc cacatctctg cctctctcat gcccccttt ctctcctgca gggtgattct  
 ggggggcctg ttgtctgcaa tggctccctg cagggactcg tgtcctgggg agattaccct  
 tgtgcccggc ccaacagacc ggggtgtctac acgaacctct gcaagttcac caagtggatc  
 caggaaaacca tccaggccaa ctccctgagtc atoccaggac tcagcacacc ggcattccca  
 30 cctgtctcag ggacagccct gacactcctt tcagacccto attccttccc agagatgttg  
 agaattgtca tctctccagc ccctgacccc atgtctcctg gactcagggg ctgcttcccc  
 cacattgggc tgaccgtgtc tctctagtgt aaccctggga acaatttcca aaactgtcca  
 gggcgggggg tgcgtctcaa tctccctggg gcactttcat cctcaagctc agggcccatc  
 35 ccttctctgc agctctgacc caaathtagt cccagaaata aactgagaag

## SEQ ID NO. 6

hk6 amino acid

40

MKKLMVVLSLIAAAWAEQNKLVHGGPCDKTSHFYQAALYTSGLHLLCGGVLIHPLWVLTAAHCKKPNLQVFL  
 GKHNLRQRESSQEQQSSVVRVIAHPDYDAASHQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWG  
 KTADGDFPDITQCAIYIHLVSREECEHAYPGQITQNMMLCAGDEKYGKDSQGDGGPLVCGDHLRGLVSWGNI  
 45 PC GSKEKPGVYTNVCRYTNWIQKTIQAK

## SEQ ID NO. 7

KLK6 nucleic acid

CDS 147.. 881

gtcgacccac gcgtccggct ggctggctcg ctctctcctg gggacacaga ggtcggcagg  
 50 cagcacacag agggacctac gggcagctgt tccttcccc gactcaagaa tccccggagg  
 cccggaggcc tgcagcagga gcggccatga agaagctgat ggtgggtgctg agtctgattg  
 ctgcagcctg ggcagaggag cagaataagt tgggtgcatgg cggaccctgc gacaagacat  
 ctcaacccta ccaagctgcc ctctacacct cgggccactt gctctgtggt ggggtcctta  
 tccatccact gtgggtcctc acagctgccc actgcaaaaa accgaatctt caggctctcc  
 55 tggggaagca taaccttcgg caaagggaga gttcccagga gcagagttct gttgtccggg  
 ctgtgatcca cctgactat gatgccgcca gccatgacca ggacatcatg ctgttgogcc  
 tggcacgccc agccaaactc tctgaactca tccagccctt tccctggag agggactgct



```

cagccaacac caccagctgc cacatcctgg gctggggcaa gacagcagat ggtgatttcc
ctgacacccat ccagtgtgca tacatccacc tgggtgtccc tgaggagtgt gagcatgcct
accctggcca gatcacccag aacatgttgt gtgctgggga tgagaagtac gggaaggatt
cctgccaggg tgattctggg ggtccgctgg tatgtggaga ccacctccga ggccttgtgt
5 catggggtaa catcccctgt ggatcaaagg agaagccagg agtctacacc aacgtctgca
gatacacgaa ctggatccaa aaaaccattc aggccaagtg accctgacat gtgacatcta
cctcccgacc taccacccca ctggctggtt ccagaacgtc tctcacctag accttgcctc
ccctcctctc ctgcccagct ctgaccctga tgcttaataa acgcagcgac gtgagggtcc
tgattctccc tggttttacc ccagctccat ccttgcata ctggggagga cgtgatgagt
10 gaggacttgg gtocctcggtc ttacccccac cactaagaga atacaggaaa atcccttcta
ggcatctcct ctcccacacc cttccacacg tttgatttct tctgcagag gccagccac
gtgtctggaa tcccagctcc gctgcttact gtcgggtgcc ccttgggatg tacctttctt
cactgcagat ttctcacctg taagatgaag ataaggatga tacagtctcc ataaggcagt
ggctgttggg aagatttaag gtttcacacc tatgacatac atggaatagc acctgggcca
15 ccatgcactc aataaagaat gaattttatt atgaaaaaaa aaaaaaaaaa aaaaaaaaaa
aggcgggccg c

```

SEQ ID NO. 8

KLK6 nucleic acid

20 mRNA join(2001..2185,3084..3135,3559..3606,4346..4502,  
8122..8369,9791..9927,11805..12483)

CDS join (3567..3606,4346..4502,8122..8369,9791..9927, 11805..11957)

```

25 acacttaaaa aatcttctga cttaaaaaaa aaagtatggt gattggaaaa tgtaaatgtg
catgcgtgct tggcatcaca ttccattggc caggacttcc ctggatgcta aaggctcctca
aatgccaggc tggggggctg ggacttggtc ccaagggaga tggggaccca gggcacgtct
gtgagaggag gggcaaggtc agcacaaggc acaggaaggt ctctctgggg caagggatac
agagaacaga gggatcctgg tccaggtggg agaggtgcag ctctgagttg gggttgaggg
30 tgtgggtaca gagaggaagg gacccccag agagaggagg cagagggata gggcctggtc
actgggttgt gcaacatcag acttgctgtc tgtgaagata gcacgtcctg agaagaagggt
gctgaggtca gtggggacca aatgtgagag ggagcaccoc gagagtatac tgaataccga
agtactcttc atccctggag tgatgggggg tgcacaatgc aagatgacaa ttagattcaa
tgcaagacaa agaaaagggt tggctgggaa cagtggctca tgcttatggt ccagctcctc
35 gggaagactg aggcgggagg gtcgcttgag cccaggaggg ttgaggctgc cagcagcaag
gatcgtgccg ctgcactcca gcctaggcga cagaacaaga ccttgtctca aaagaaaaaa
gaactttttt ttttaagtta cctgtagtgc ccagcccaag cagggtgctga gccagacttc
attcctatca ttgtccttat tacgcagtga cttccccctc ctcatctctc tccactctgc
cacgcacaca ccctcacctc ccagcccata ccaaccaccc caaccactgc ctgtggtttc
40 ccatgtgcac ccaggccagg cattttcacg gcctttcctc ctgacctacg cctggctcag
ctttctaggg ccaagttaaa agacacctcc ctaaaatctc ccagatccct ctgctactgc
ccagcaccac catcttatca cagccccacg tcgttcccaa gtgctctccg atttctgctt
aactccatgc ctctcgctgt gtgtccgcat ctcatcaata agtcctcaag tcctcttcca
tcttgctagc ttctcctcgt ctcggggaatc atccccgcta ctctctgggg aaactgactc
45 ccttctgggc acacacagtg ctacccccgg ggaaatctaa gaagagaccc aggagaagat
aagcacggag agtcagagaa tcaaggggaa agaaaggag agaggccggg cacagtggct
cacacctgta atccagcact ttgggaggcc aaggtgggtg gatcacctga ggtcaggagt
tgagaccag cctggccaac atggtgaaac ctcttcccta ctaaaaatac aaaaaacatt
tagccggggc tgggtggtgg tgctgtaat ccagctact tgggaagctg aggcaggaga
50 actgcttgaa ctgaggaggc ggaggttgca gtgaactgag atcacaccac tgcactccag
cctgagtgac agagcaagac tccgtcaaaa aaaaagaaag aaagaaagaa aagaaggaaag
gaaagaaaga aggaaggaaag gaagggagga agggagagag gaagggagag aggaaggagag
agagagaaaa aaagaggag agagacacaa atacagagac tgagatggga gagagagaga
gatggaagct ccctcccctc catggccagg gagacagatg gagcaagaga cctcaggggt
55 gggcagactt ggaggagaag gaccaggagg atgtggagtg ccgaaatctc cagtcagggc
caggtgggca gtcagagact gcaaaggagg actgtcagac agggacaaaa ggaagccatt
gatgtaaccg ccctcccgcc tgcccgccgg aagagagggt gaggccggag ctgctgggag
catggcactg ggggtgctggg aggcggacaa agcccagatt ttcctgggcc ctttccccat

```

5 cgcgcctggg cctgctcccc agccccgggc aggggcgggg gccagtgtgg tgacacacgc  
 tgtagctgtc tccccggctg gctggctcgc tctctcctgg ggacacagag gtcggcaggc  
 agcacacaga gggacctacg ggcagggtgt tgagtcaccc caaccgcact gaacctgggc  
 aggtgcttc ccagtgccgg agggctctag agcccgaggt gagggcctgc aggtccctgg  
 gtggcacaga gagtgtctgg ggtgcaggga ggcctggggc accatctgct tgccccagag  
 gccggaatth gtcttcagac actttctttc tccaaaaccc ggaggtctaa ggactgagcc  
 gactagaact tctctgtcct cagattcagg cccagcccc tcttccctca gaccacggag  
 tttaggctcct agccccctcct ccttcagacc caggagtcca agttcccacc tcttccctca  
 gactcaggag tccaggcccc cagccctcct tccctcagac ccaggagtcc aagttctcac  
 10 ctctctccctc agaccagga gtccaggccc caagcccctc ctccctcaga cgcaagggtc  
 caggccccca gccctcctc cctcagactc aggagtcag gcccccgaag ccttccctcc  
 tcagaccag gagtccaggc cctcactgca ctccaggacc agtgcctcct tccctggagg  
 cctggtcagg ggtcaccaag agcagagcgt gggggcggga ggaatgtgtg tgggaggcct  
 gggttaaggag gaaaagggtg tagccagctc cctggctcag ggacctgaga gacagggtt  
 15 aaaaggacgt tccagaagca tctggggaca gaaccagcct ctccaggga ggcctgggag  
 ctgggggtgt gtgtctggca gtccctgcag ccctgggctc tgcggccctc gcgtcctccg  
 cttggtctct ccactgcac tgagtgtctt ctctcctcac ggctccccgc atttctaact  
 cttctgtcct cctcgtctca aagctgttcc ttccccgcac tcaagaatcc ccggaggccc  
 ggaggcctgc agcaggtgag atcacagaca gcacagaacc tgcgggtgtg gcgggtgtgg  
 20 tggccattgc gcacagagcc aggtccgag gaaaactccc atacagagga agaacgtag  
 ggccccctat ggtaacctc tctgtcgcac aggaaggcaa atcagtgcct aagaaagtag  
 aaagatctaa tcagaatctc accatgggtt actggaccag tggacgtagt tgaattctct  
 ttggcactgt tttcgtggat cctcttgaa gatgtgggt gaggaagaat aaatcaggag  
 gctagatggg aaggacagag gtcaaggcag gagaccatag caggccagga aggaaggaga  
 25 ggatgcagag ggagcagaca gagggtggg gggagggtcg aggcagtgc taatggacca  
 tgtggcttcc cctctcagga gcggccatga agaagctgat ggtggtgctg agtctgattg  
 ctgcagggtg ggaaagggca tttggatggg ggaggcttg agacagggtt gggcttgttg  
 atggagaaga ggctggtatt ggggatggg atatgcacag ggttgggtg ggggagcttt  
 gaaatgagga agacgttgg gattaggcta aggggtggga atacagatag ggagggtggt  
 30 gggagggtgg tttgaagata tgagggtttt ggggtgggtt ggctttaggg atggggatct  
 aaacatagaa gaggtaggag gttaggttga aagttggaga gagcccgga ataggggata  
 cagttgggtt tgtaatggga atgggttaag tttgggagtg gaaatacaga gaagcttttt  
 ttttttgaga cagggtctca ctctgtcacc caggctggag tgtagtgga tgatccatag  
 35 ttcaactgcag acttgaactc ttgggtctca agtgaccctc ccacctcagc ctcccaagta  
 gctgggacta caggcgtatg ccaccatacc ctgctaattt gtgtgtgtgt gtgtgtgtgt  
 gtgtgtgtgt gtgtgtgtgg agatgaggtc tctactgtgt accgaggctg gtctcaaaact  
 cctgggtcctc agcgatcctc ctgcctcagc tgggattaca ggcataagcc actgcacctg  
 accaatcttg actggagtcc atgttgaggg ggtgcgctt ggtttctcca gaactcctct  
 40 ctgactcaga tcttctctcc ctcagcctgg gcagaggagc agaataagtt ggtgcattgg  
 ggaccctgcg acaagacatc tcaccctac caagctgccc tctacacctc gggccacttg  
 ctctgtgggt gggctcctat ccatccactg tgggtcctca cagctgcccc ctgcaaaaaa  
 ccgtgagtct acactgtaaa tgaacagcag atgcgactga accctgaggg tgtcttatag  
 atgtcaggca ggaggtgaca taggcattccc ccccatccca gcacgaggcc atctgatagc  
 45 cagggtgcatt cggctgttg ttaattgagt acttaattgt tgccaggccc tgcgggcata  
 gcagtggaaa agaaaataaa aaaaagaaaa caaaaaaaa caagcaaat tgctgttttc  
 ctgaacttac tttctaatgg gggaattgga tcatttgggg acctgcaggg cgtgatgggc  
 atttggattt aattctgagc acagtaggaa gccactgggc agttttgttt ttgtgtttg  
 tttgttttt gagacacagt ctgcctctgt caccaggtt ccagcgattc tctgcctca gcaccccaag  
 50 cagctcactg caacctctgc ctcccaggtt cctggctaatt tttgtatgt ttggtagaga  
 tagctgagat tacaggtgtg caccaccttg cctggctaatt tttgtatgt ttggtagaga  
 cggggtttca ccatgttggc caggctggtc tcgaactcct gacctcaggt gatccgccc  
 cctcgccctc ccaaagagct gggattacag gcatgagcca ccaccacacc cagcctgatt  
 tacattttta caagcacctt ggctaccacg tggaaactgg totgggcaag agagaggag  
 55 ggaggccac gtgggggctg ttgctttcat ccggcgacat aggggggtgg cttgaacca  
 ggggtcgca gtggggatgg agggatgttg aatatcttgg gatgtggaat tctgagactg  
 agccagcaga atctggcaac gaggaacagg agggagagga agaagcacgg ctggcttccg  
 tgtattttgt ctgaacaact ggggtttttg ccacgtcttt ctctgagttg tgggagaggg  
 aaagagaaac aggcggggtg taggcagggg agcatctgac attttgcttt agccacgatg  
 agttggagat gccggggaga tgtcccagca gggaggccag ggaggactct ggagctcaga

	ggagagggtca	gggctggagg	ttaaaatgaa	ggcatcgtca	gcaaacaggt	gtattttaaag
	ccatgggact	agatgagatc	atccaaaaag	ctggcatagt	tggaggagct	ggaggggccca
	ggacaaaaac	cctgggcgct	gatcctcact	agtcagattc	acgacagctg	ccacttgttt
	gatgtctaact	accaatcagg	tgctgagtga	aacctatgtac	acacctttcc	tggaaatgcc
5	accacaagg	actcttggca	ccattttgca	aatgaggaaa	ctgaggtgca	gggaaatagc
	aagtgcacaa	ccctgggggtg	gttcccctga	ccccaaggag	accttggatg	actctcacca
	ccatcattca	ttcctttgat	gtacattgac	taagagcacc	tgctaagtgc	cacattcgag
	ttgggcagtg	gagattcagc	aatggatggg	acacacacgt	catccctgcc	ctcgggagca
	caaggacaga	aaggtgcaga	caagcaaaagt	gagggctggg	catggtggct	cacgctgtta
10	atcccagcac	tttgggaggg	cgagggtggg	ggattacctg	agttcgagac	cagcttggcc
	aacatggctc	aaccctgtct	ctactgaaaa	tacaaaaaat	tagccaggcg	tgggtggtgg
	cttctgtaat	tccagcaact	tgaggagcta	aggcaggaga	attgcttgaa	cgtgggaggg
	ggaggttgca	gtgagccgag	atcgcgccac	tgactccag	cctgaaccac	agagcgagac
	tctgtctaaa	aaaaaaaaaa	ggaaagaaaag	aagcagcaaa	ttgggctggc	cgtggtggct
15	catgcctgta	atcccagcac	tttgggaggg	cgaggcgggg	ggatcactcg	agcccaggag
	tacaaagctg	cagtgcagctg	tgatctacag	aacaccactg	cagatccagc	ctgggtgaca
	gagcgagacc	ctgtctcaaa	aaaacaaaca	aacaaaagaa	gcaaaccctt	caaaacccca
	tataattaca	aattatgaag	gaaaagaata	cgggtaccta	ctttagatgg	aggagggtca
	ggaaggactt	tctaattgaga	taaaatccaa	gcggaggcat	gaagatggga	aaaggaattg
20	tcagggcaga	ggaaaggctg	tgataacacc	cctgaggtga	gaaccgtctt	gagtattctc
	agaaaataaa	atttcccgtt	cactgggggg	cagaagggtgc	tgggagataa	ggttggaaag
	tgactacagc	cagatcacac	aggggctoca	gtgccaagtg	gaggagccca	ggctttattc
	ttaggacaa	ggggagccat	gggtgatgtc	tgagcaaggg	agtgactctc	tgtttcagga
	atatgtatca	aacacctatc	ctgtgcoagg	tgctgatcaa	cgactggag	atactatatc
25	tgaatagaac	aaaaatcccc	atcttgacat	cttagagctg	cactgtctaa	tatggtagcc
	atcagccaca	tatagcaaat	tacattgaaa	ttaatgaaat	ggaaaatcca	caagccacat
	ttcaagtact	cagcagccac	ctgtagcttg	tggttccccc	agccacctct	ggacagtga
	gatcgagatc	atggcatcgt	agcatttagt	ggacagcatt	gctctgcaag	gaggagaaat
	aacacaatga	gtaaatattt	aacaataaat	atatagoagg	tcggatgatt	gtgatagggt
30	ctctggtgga	acagaaagca	ggggaggggg	ataggaattg	cctactaaca	ggtatttgta
	tttttaattgg	gcaactaagg	aaggcttccc	tgagaggcga	cattttaaagg	aagtgaggga
	gtgagctatg	cagatacttg	gaggacagac	ttgctggcag	agggaaacagc	agtgcacagg
	ccctgaggtg	ggaagatcac	tattgtgttc	aaggcaagac	agggaaagcca	gcgtttggct
	ggagcagagg	gagagaagg	gagagtggga	ggagaagatg	tctgtgagat	gatggggcag
35	tgcttgcaag	gcctggtgtg	ccacgttgag	aactttggct	ttgattctga	gtgagatggg
	agtcatagga	ggggctgagc	agaggaggca	caggaccaac	ttacattggt	aaaatatctc
	tgggtgcttt	gtggaggatg	gactgtgggg	gaccagagac	agagcaggga	gcccagtgag
	gaggctactg	ctctagttca	ggtaggaagt	gaaaaggcag	ctcaaaccaa	gatggtagcc
	gtgggaaagg	tgagatgtgg	ccagattctg	gatatgcttc	agagaggcaa	aaggaattct
40	ggacagcttg	gatgtagggc	atgaaataaa	gagagtgaag	aatagccccc	aagattattc
	tgaaaggatg	gaattgccat	ttaccagct	ggggaagact	gtgggaggag	caggccagcg
	attcatgact	tcccagccct	ctctgaagcc	tcaactgcag	cccaagggct	ccaggtgaga
	cccagccctc	ttccttccca	ggaatcttca	ggtcttcctg	gggaagcata	accttcggca
	aaggggagagt	tcccaggagc	agagttctgt	tgtccgggct	gtgatccacc	ctgactatga
45	tgccgcccagc	catgaccagg	acatcatgct	gttgcgcctg	gcacgcccag	ccaaactctc
	tgaactcatc	cagcccttcc	ccctggagag	ggactgctca	gccaacacca	ccagctgccca
	catcctgggc	tggggcaaga	cagcagatgg	tcagtagtgg	gaggctgggtg	gggagcaggc
	tactggtctac	ttgggggaagt	gtgccaagg	atggggagtg	ggaaaattgg	tgagggggcca
	tgggaagatg	ggctaattggt	gaggaccaat	gggacagggt	ttcaatggga	gaaagggtcaa
50	gggggagggg	gagtgaattt	gggagctggg	ccagtgaagt	aacagccaat	ggaaaatgta
	gaccaatggg	tgaatagcat	gggagagatg	gaacataaga	tgaagggtca	ataaagaggg
	aaggtcagtg	gggagatgct	aatcagggaag	gatgtcaaag	gtcaaagggg	actgatcagg
	attcattgaa	cagcaggaag	gaataatgga	gaagggaactg	atggaagaag	agaaccaaat
	aaagcacaaa	agccaactga	aggatgtgaa	ttgagacagt	gaatgggggt	atagctgatg
55	tgaaggggac	taaggggaaa	ggatcaatgg	tccagaggag	tcactagagg	aaaaaacagg
	tccaatagat	cagcaggatc	catgaagggtg	ggcctgtgtg	tgaaggggcca	ataagaaagg
	tgaaccattg	gatgaagggc	cagtgggaag	gcagagacaa	tgggggagga	tgccggcaagt
	tagaaaagga	ccaatgaggg	aggtggacca	ttggatgaag	ggctaatagg	aaggggagagc
	cagtggggga	tggtgaggcc	agtttagaaa	ggaccaagga	gggaagcaga	ccaataggaa

gagagagcca atgagggagg gcagggccag ttaggaaagg accaatgagg aaggtagacc  
 attggaggaa gggccaatag aaagggagga tccatgagg aggggtggga cagttagaaa  
 aggaccaatg atggagggtg accattggat gaagaaccaa tagaaaggaa gaaccaatgg  
 gagagggcat ggccagttag gaaaagacca atggtcacag agtgaccaat caagatgaat  
 5 caatgggcag gaagtgtcca atgaagaatg gactactgat caggaggggt acagtagagg  
 agggcgtaac agaggaagag tcctccaggt caactgaaac tactgaagaa ggtgggacca  
 gtggaagaga gaaaagtgga ggagggaacct aagagaaaaag gaaaaccaat aggaaatgag  
 gactcctgga gaagagacta ttaatgagga agacagccaa tgggggggaa gaatgataga  
 aagagggacc aattaggagg cagggacgat ggtaatgaga tgtaagaatg agagacaaac  
 10 aggaagagggt gtgccaatag aaaagaggga ccaatagagg atggaggact tataggggtt  
 ggggggtgac tggggagggt gaggggagtg caaggcctgg gctgagtctg gcccatctct  
 cccctaacag gtgatttccc tgacaccatc cagtgtgcat acatccacct ggtgtcccg  
 gaggtgtgtg agcatgccta ccctggccag atcaccaga acatgttgtg tgctggggat  
 gagaagtacg ggaaggattc ctgccaggtg aggtgaccog gatctgccac ttacacagcc  
 15 agggacagga cgaagtcaca aaaacatggc cagacacagg aagagagaga cacaggccaa  
 aagagagctt tacagagaca gatagagaca ggctgaggga gaaccaagc ctgaaaaga  
 agagacttag ttcaacacac agagacacag tcagggatata gcagagatat aaagacacag  
 ccagacagaga caggaagtgc agagacaagg atggaggccg cggtatcaag aaccagagag  
 gccaggagca gcggtcatg cctgtaatcc cggcactttg ggaggccgaa gcaggaggat  
 20 cacctagggt caggagtctg agaccagcct gatcaacatg gtgaaacctt atctctacta  
 aaaatacaaa aattaggatg ggcacagtgg ctcatgcctg taatcccagc accttgggag  
 gccgaagcag gaggatcacc tggggtcagg agttcgagac cagcctgagc aacatgggtga  
 aacctatct ctactaaaaa taaaaaatt aggatgggca cagtggctca tgctgtaat  
 cccagcacct tgggaggccg aagcaggagg atcacctggg gtcaagagat tgagaccagc  
 25 ctggccgata tggtgaaacc ctatctctac taaaaataca aaaattagct gggcctggtg  
 caggcgccctg tagtcccagc tactcaggag gctgtggcag gagaatcact tgaacctgga  
 ggcggagggt gttgcagtga gtcgagatca tgctactgca ctccagcctg gcaacagagc  
 aagattccgt ctcaaaaaaa aacaaaaaa ctggagaatt gcttaaaccc aggaggcaga  
 ctgtagtccc agctactcgg gaggtgagg actccagcct ggggacagag ccagactctg  
 30 ggctgcagtg agctgagatc acgccactgc gtaaggaagc agatgggtgtg aggggactgt  
 tctaaaaaca aaaagaacca aagagaagta gtcagaaac gaccaggctc tggaggagggt  
 ccttcctcaa acagagcccc ggactccctc ccgaatactt taacttgggt ttccctcatt  
 agacactagc tggggaaaag tctttatgcc agaatactta atgcactgag ggaatgtcag  
 35 agaccaacca agggggagac acaggcagaa acggagacac aggcagaaac agggacagag  
 acagggaaag cgatacatag caagtggac gcaaagaaag ggcaggtggg cgagactgtc  
 ctcaagacac gaggtggaga ggtgtccctg gacagaatag tgccaggcat atctctccct  
 gggccctccc tacctctccc acctgggtct tatcgtctcc tccctccccct cctccctctc  
 40 ctctctctcc tctctctctc cctccctcat catcttcttc tttctctctc ctctccatcg  
 gtctctacac ctctgcctct ctccacacct ctcaagtctcc attcttaaat tgtttctctt  
 tctgtctctc tatgttctct tgcatcttgg cattcctatc tctgtgtctt tgagtctcct  
 ttattctctc tctaccattc tctctctgtg cctttgtgtg tcttactgtc tctctctctg  
 tctctctgtc cctgagctct tctctccatc tttcagtaag tacctctgtc cctttctacc  
 45 tctctctctg tcacacacac acacacacac acacacacac acacacacac acacacagtc  
 tctgggtttc tatctgtatc tgactttctc cctctttcct gcagggtgat tctgggggtc  
 cgctgggtatg tggagaccac ctccgaggcc ttgtgtcatg gggtaacatc cctgtggat  
 caaaggagaa gccaggagtc tacaccaacg tctgcagata caggaactgg atccaaaaa  
 ccattcaggc caagtgacct tgacatgtga catctacctc ccgacctacc accccactgg  
 50 ctggttccag aacgtctctc acctagacct tgcctccctc cctctctgac ccagctctga  
 ccctgatgct taataaacgc agcgacgtga gggctcctgat tctccctggg tttaaccag  
 ctccatcctt gcactactgg ggaggacgtg atgagttagg acttgggtcc tcggtcttac  
 cccaccact aagagaatac aggaaaatcc cttctaggca tctctctcc ccaaccttc  
 cacacgtttg atttcttctc gcagaggccc agccacgtgt ctggaatccc agctccgctg  
 55 ctactgtcg gtgtcccctt gggatgtacc tttcttctact gcagatttct cacctgtaag  
 atgaagataa ggatgataca gtctccataa ggcagtggct gttggaaaga ttaaggttt  
 cacacctatg acatacatgg aatagcacct gggccaccat gcactcaata aagaatgaat  
 ttt

## SEQ ID NO. 9

KLK6 nucleic acid

5 CDS 246..980

```
aggcggacaa agcccgattg ttcctgggccc ctttcccat ogcgccctggg cctgctcccc
agccccggggc agggcgggg gccagtgtgg tgacacacgc tgtagctgtc tccccggctg
gctggctcgc tctctcctgg ggacacagag gtcggcaggc agcacacaga gggacctacg
10 ggcagctgtt ccttccccg actcaagaat ccccgaggc ccggaggcct gcagcaggag
cggccatgaa gaagctgatg gtggtgctga gtctgattgc tgcagcctgg gcagaggagc
agaataagtt ggtgcatggc ggaccctgcg acaagacatc tcacccttac caagctgccc
tctacacctc gggccacttg ctctgtggtg gggtccttat ccatccactg tgggtcctca
cagctgcccc ctgcaaaaaa ccgaatcttc aggtcttcct ggggaagcat aaccttcggc
15 aaagggagag ttcccaggag cagagtctctg ttgtccgggc tgtgatccac cctgactatg
atgccgccag ccatgaccag gacatcatgc tgttgcgcct ggcaagccca gccaaactct
ctgaactcat ccagcccctt cccctggaga gggactgtc agccaacacc accagctgcc
acatcctggg ctggggcaag acagcagatg gtgatttccc tgacaccatc cagtgtgcat
acatccacct ggtgtcccgat gaggagtgtg agcatgccta ccctggccag atcaccacga
20 acatgttgtg tgctggggat gagaagtacg ggaaggattc ctgccagggt gattctgggg
gtccgctggt atgtggagac cacctccgag gccttgtgtc atggggtaac atccccctgtg
gatcaaagga gaagccagga gtctacacca acgtctgcag atacacgaac tggatccaaa
aaaccattca ggccaagtga ccctgacatg tgacatctac ctcccgacct accacccac
tggctggttc cagaacgtct ctcacctaga ccttgcctcc cctcctctcc tggccagctc
25 tgaccctgat gcttaataaa cgcagcgacg tgagggtcct gattctccct gggtttaccc
cagctccatc cttgcatcac tggggaggac gtgatgagtg aggacttggg tcctcggtct
tccccccacc actaagagaa tacaggaaaa tcccttctag gcatctctc tccccaaacc
ttccacacgt ttgatttctt cctgcagagg cccagccacg tgtctggaat cccagctccg
ctgcttactg tcggtgtccc cttgggatgt acctttcttc actgcagatt tctcacctgt
30 aagatgaaga taaggatgat acagtctcca tcaggcagtg gctgttgga agatttaaga
tttcacacct atgacataca tgggatagca cctgggccc catgcaactca ataaagaatg
tatttt
```

35

## SEQ ID NO. 10

hk7 amino acid

```
MARSLLLPLQILLLSLALETAGEEAAQGDKIIDGAPCARGSHPWQ
40 VALLSGNQLHCGGLVNERWVLTAAHCKMNEYTVHLGSDTLGDRRAQRIKASKSFRHP
GYSTQTHVNDLMLVKLNSQARLSSMVKKVRLPSRCEPPGTTCTVSGWGTTTSPDVTFF
SDLMCVDVKLISPQDCTKVYKDLENSMLCAGIPDSKKNACNGDSGGPLVCRGTLQGL
VSWGTFPCGQPNDFGVYTQVCKFTKWINDTMKKHR
```

45

## SEQ ID NO. 11

KLK7 nucleic acid

CDS 16..777

```
ggatttccgg gctccatggc aagatccctt ctctgcccc tgcagatcct actgctatcc
ttagccttgg aaactgcagg agaagaagcc cagggtgaca agattattga tggcgcccca
tgtgcaagag gctcccaccc atggcagggt gccctgtcga gtggcaatca gctccactgc
ggaggcgtcc tggtaaatga gcgctgggtg ctactgccg cccactgcaa gatgaatgag
tacaccgtgc acctgggcag tgatacgctg ggcgacagga gagctcagag gatcaaggcc
55 tcgaagtcat tccgccaccc cggctactcc acacagaccc atgttaatga cctcatgtc
```

```

gtgaagctca atagccaggc caggctgtca tccatggtga agaaagtcag gctgccctcc
cgctgcgaac cccctggaac cacctgtact gtctccggct ggggcactac cagcagccca
gatgtgacct ttccctctga cctcatgtgc gtggatgtca agctcatctc ccccaggac
tgcacgaagg ttacaagga cttactggaa aattccatgc tgtgcgctgg catccccgac
5 tccaagaaaa acgcctgcaa tggtgactca gggggaccgt tgggtgtcag aggtaccctg
caaggtctgg tgtcctgggg aactttccct tgcggccaac ccaatgaccc aggagtctac
actcaagtgt gcaagttcac caagtggata aatgacacca tgaaaaagca tcgctaacgc
cacactgagt taattaactg tgtgcttcca acagaaaatg cacaggagtg aggacgccga
tgacctatga agtcaaattt gactttacct ttccctaaaag atatatattaa acctcatgcc
10 ctggtgataa accaatcaaa ttggtaaaga cctaaaacca aaacaataa agaaacacaa
aaccctcaa

```

## SEQ ID NO. 12

## KLK7 nucleic acid

15

mRNA

```

join(1756..1785,3179..3309,3722..3869,4566..4813,5129..5265,7362..8265)
/product="stratum corneum chymotryptic enzyme" /note="alternatively
spliced"

```

20

```

mRNA join(1756..1785,3179..3309,3722..3869,4566..4813,
5129..5265,7362..7991) /note="alternatively spliced"

```

mRNA

25

```

join(1821..1864,3179..3309,3722..3869,4566..4813,5129..5265,7362..8265)
/product="stratum corneum chymotryptic enzyme" /note="alternatively
spliced"

```

mRNA

30

```

join(1821..1864,3179..3309,3722..3869,4566..4813,5129..5265,7362..7991)
/note="alternatively spliced"

```

```

CDS join(3237..3309,3722..3869,4566..4813,5129..5265, 7362..7517)

```

35

```

ggcatggtgg tgcacgcctg taatccagct actcaggact ctgaggcagg agaatcactt
gaacacgggg gactggagggt tgcagtgagc cgagatcgtg ccattgcact ccagcctggg
tgacagagcc agagtccatc aaaaaaaaaa aaaaaataaga aagattcttc tctcctctat
gtgtccatgc agtctcatca tttagctacc acttgtaagt aggaacatgc catatctggt
40 tttctgttcc tgccttagtt tgtaagggtg atggcctcca gctccattca cgtccctaca
aaggacatga tcgtgttctt ttttatggct acgtagtatt caattgtgta tacgtaccac
atcttcttaa tccagtctat cactgatgga catttaggtt gattccctgt gtttgcctgt
gtcaatagtt ctacaatgaa cgtacgtgtc catgtgtctt taaacagaat gatttatatt
cctttgggta cacacactgg ggcttatgag aggggtggaga gtgggaggaa ggagaggatc
45 agaaaaaaat aactaatggg tactaggcct aatacctggg tgattaaata atctgtataa
caaaccoccc tggcgcacgt tcacctacgc aacaaacctg cacatcctgc acatgtaccc
ccgaactgaa aagttaaaaa aagaaaaata aatatttgct tataaattaa taaatgaagc
cctcaaaaaa gttctattag ataattgtta gtacagacat ttttgttata aatacataat
atacaaagaa atctatgtat aacatgatta aatgaccat aagaacatag atcctaaaca
50 tggcaaatat tagtgggggtg gggttaggga aagcgttgtt ttttaacttac acctctctgt
tagagtggg aatgggttca ggcgtaatta caggcacgac tgggatcagc ttggacaagt
tccccaggc gggccagaat taggatgtag ggtctaggcc acccctgaga gggggtgagg
gcaagaaaaa gggcccagaa gccgggagca gtggctcagc cctgtaatcc cagcactttg
cggggcccag gcgggcacat catgagggtc ggagatcgag accattctgg ccaacatagt
55 gaaacccggt ctctactaaa aatacaaaaa ttagctggga gtggtgggtg gtgacctgtaa
tcccaggtag tcgggaggct gaggcaggag aatcacttga acctgggagg cggagctggc
agtgaagcca gatcgcgcca ccgcactcca gcctggcgat agagagagac tccatccaaa
aaaaagaaag gaaggaggagg agggaggagg gaagaaagaa agaaaaccgc ccagagagaag

```

	gacccgagcc	agagcctatt	ctctgagctc	agcgactgct	tgaatcccgc	tcctgcccct
	cagacccagc	gcacccgggtc	cctcccccca	gagcagccag	gagggactgt	gggaccagaa
	tgtgcggggg	cgcaggagct	gggcaccgcc	cgctccttcg	agggagggtg	gagagagagt
	gcagtggtgc	caattgtctc	cgctgcgtca	gggttccaga	taaccagaac	cgcaaatgca
5	ggcgggggtg	tcccagagtc	ggctccgcct	gcacccaggg	gcgctggggc	cgggcatggg
	gcggggggtg	atataagagg	acggcccagc	agagggatga	agattttgga	gcccagctgt
	gtgccagccc	aagtccgaac	ttggatcaca	tcagatcctc	tcgaggtgag	aagaggcttc
	atcaagggtg	cacctgtagg	ggagagggtg	atgctggctc	caagcctgac	tctgctctcg
	agaggtaggg	gctgcagcct	agactcccgg	tcctgagcag	tgagggcctg	gaagtctgca
10	atttgggggc	ttttagggaa	aaacgaacta	cagagtcaga	agtttgggtt	ccacagggaa
	gggcaagatc	ggagcctaga	ttcctgggtc	tctagggatc	tgaagaacag	gaattttggg
	tctgagggag	gaggggctgg	ggttctggac	tcctgggtct	gagggaggag	ggcctggggg
	cctggactcc	tgggtctgag	ggaggagggg	ctgggggtct	cgactcctgg	gtctgaggga
	ggaggggctg	ggggcctgga	ctcctgggtc	tgagggagga	ggggctggga	cctggactcc
15	taggtctgag	ggaggaggag	ctggggcctg	gactcctggg	tctgagggag	gaggggctgg
	ggcctggact	cctgggtctg	agggaggagg	ggctggggcc	tggactcctg	ggtctgaggg
	aggaggggct	ggggcctgga	ctcctgggtc	tgagggagga	ggggctgggg	cctggactcc
	tgggtctgag	ggaggagggg	ctggggcctg	gactcctggg	cctgagggag	gagggactga
	gacctggact	cctaggtctg	agggaggagg	gactgggacc	tggactcctg	ggtctgaggg
20	aggaggagct	gggggctggg	actcctgggt	ctgagggagg	aggggctggg	gcctggactc
	ctgggtctga	gggaggaggg	gttggggcct	ggactcctga	gcctgaggga	ggagggactt
	ggacctggac	tcctaggtct	gagggaggag	gagctggggg	cctggactcc	taggtctgag
	ggaggcgggg	ctgggggctc	ggactcctgg	gtctgaggga	ggagggggtg	gggocctggac
	tcctgagcct	gagggaggag	ggacttggac	ctggactcct	aggtctgagg	gaggaggagc
25	tgggggctg	gactcctagg	tctgaggagg	gaggggctgg	gggcctggac	tcctgggtct
	gagggaggaa	ggtgctaggg	tctggactct	tgggtatgag	ggaggaggag	gttaggggtc
	tggacttctg	agtgtaaagg	aggagaggcc	agagaaagga	atttctgggt	ctgaggggag
	aggggctggg	gttctggacc	cctaggtctg	agggaggagg	ggctggggcc	tggacccctg
	ggtctgaggg	aggaggggct	ggggccggta	ctcctgggtc	tgtgggggga	ggggctgggg
30	cctggacccc	tgggtctgag	tggggagggg	ctggggcctg	atgctttctc	cttctcagct
	ccagcaggag	aggcccttcc	tcgctgggca	gcccctgagc	ggctcagcag	ggcaccatgg
	caagatccct	tctcctgccc	ctgcagatct	tactgctatc	cctagccttg	gaaactgcag
	gagaagaagg	tgaagactgg	actgggaagt	ctgacctcac	ctcagggccc	ccactgaccc
	tctccaagga	gtccctgagt	cagaaccctt	ccctcctcaa	acagcttcca	tcctgggagg
35	accagactgt	cggctgaagc	ccccgctott	cctgcttctg	ctgactcagg	gggtctctgt
	cccctccagg	ccctgcctcc	tgtgctcagg	gtctctctgt	ggttccccag	atgagatgcg
	cctcctgggt	ttctgagtgg	gctccttctg	tctgtctcta	tcctatctct	ttgctttctc
	tgtattttct	cacacatttt	catctgtctc	tgtccatctc	tgactctggg	aatccctgag
	gtgcagcctc	agccttcccc	taatgctagc	tacccaagtg	ctcctccatg	tctccatcca
40	gcccagggtg	acaagattat	tgatggcgcc	ccatgtgcaa	gaggctccca	cccatggcag
	gtggccctgc	tcagtggcaa	tcagctccac	tgcgaggcgg	tcctgggtcaa	tgagcgctgg
	gtgctcactg	ccgccactg	caagatgaag	taggtgccgc	ccaagtctct	gctggagggtg
	caccagcgct	tcagctcgcc	tatgggggtg	gaagggcagt	ccttctgtgc	ctacggctct
	attctcctct	ctctgggtct	ctgtcctcct	ctctctgggc	ctctgtaccc	cctctccctg
45	gggctctgtc	ccctctctct	cctggctctc	tgtctccctc	tctctgggtc	tctgtccccc
	tctctctgga	tctctgttcc	cctctctctg	tgtctctgtt	ccccattctc	tctaggtctc
	tgttccccct	cctctctctc	tgggtctctg	tccctctctc	tctggatctc	tgtcccccct
	tctctctggg	tctctgttcc	cctctctctg	ggtctctgtc	ccctctcctc	tctctgtgtc
	tctctccccc	ctcctctctc	tgtgtctctg	tccccctccc	tatctctgtg	tctctctccc
50	ccctcctctc	tctgggtctc	tgtccccccc	tctctgggtc	tctgtctccc	tctctctggg
	gctctgtccc	cctctctctc	tggatctctg	ttccccctct	tctgggtctc	tgtctccccct
	cctctctctg	tgtctctgtc	ccccctcctc	tctctgggtc	tctgtcccca	ccccgtcccc
	caggtctttg	cacacccctc	ctgtcacagt	gtctcttctg	aatctgtgaa	tgtcactcct
	cgcagtgagt	acacccgtga	cctgggcagt	gatacgctgg	gcgacaggag	agctcagagg
55	atcaaggcct	cgaagtcatt	ccgccacccc	ggctactcca	cacagacca	tgttaatgac
	ctcatgctcg	tgaagctcaa	tagccaggcc	aggctgtcat	ccatgggtga	gaaagtcagg
	ctgccctccc	gctgcgaacc	ccctggaacc	acctgtactg	tctccggctg	gggcactacc
	acgagcccag	atggtaggtg	gcctcagtga	cccaggagtg	caggccccag	ccctcctccc
	tcagaccag	gagtccaggc	ccccagcccc	tcctccctca	gaccaggag	tcaggcctc

	agccccctct	ccctcagacc	caggagtcca	ggccccagc	ccctcctccc	tcagaccchg
	gagtcagac	cccagcccct	cctccctcag	accagcagct	cctggggccc	agaccctcct
	ccctcggaac	caggagcctg	aacaacagcc	cttctgggtcc	tcgcccccat	cctctctgac
	tgacagctct	ccctgctcct	ccctgcagtg	acctttccct	ctgacctcat	gtgctggtat
5	gtcaagctca	tctcccccca	ggactgcacg	aagggtttaca	aggacttact	ggaaaattcc
	atgctgtgcg	ctggcatccc	cgactccaag	aaaaacgcct	gcaatgtgag	accctccccc
	ccaattcctc	cccagtcctg	ggtaccctgt	ctgcatgccc	cagggacaga	gcttgaccgg
	agtgaactggg	taccaagccc	ggccttgccc	tccccccagg	cctggcctcc	tcagcttttt
	ccacctcatt	ctctgcctag	gtcaggggtg	ggagtttact	taggggccaa	tgtggccctg
10	gggatgagac	agagagttta	ataggggtga	gaaagtgggg	gtgggaccag	ggaaggagac
	tgagggtgctg	gcctcaggcc	caaaccttaa	gggggcacca	aaaacctcag	tgattgagat
	aaatcataat	gcaatattta	aaaataaaaa	taaaaactca	tcagaaagtc	catgatggac
	aaaatgtcac	attttaaata	aagagcaggt	ggatcttact	gaattttccc	ttgcccgaag
	tactagcgtg	gtcagcaca	gcgctgtact	ggcactgtct	tcatttaaaa	tgtggatacc
15	atgcccatca	tgagttttta	tgtattacat	ttgatttcgt	taagtactgc	attgaagtat
	tgtgtattgc	agtactgag	attttgtgcc	tgaagctgat	gactcactca	cctgaccctg
	gccctgggtcc	cggggaaaac	actctttctc	tccaactcct	ctctgttccc	tctttctggc
	cttttgcctc	ccctctgtt	tctgaacagt	cttccacat	ctctctttgt	gacataattt
	catttcattc	ttttcctctt	tgttttttct	ctgtgttgag	ctagcttgct	ctccctccct
20	tgttctctct	ccatgccctc	ctctctgtct	tctgtcttct	ccctctttct	cttgtctctc
	tctctctcct	ccctccctc	tctcctctcc	ctgccccctc	gctctctctt	ttttcctctc
	tctctgtctc	ctctctggcc	ctctcctctt	tctctctctc	ccccacttct	ctgtctctct
	tcactctctc	ccctcatctc	tccttgcccc	ctccttttta	ctgtctctct	ctttctcttt
	cttctatctc	tctcctctcc	ctgcccctcc	cccatctctg	tctttctttc	tctctcttta
25	ttctctctct	ctcttccagt	ctctctctcc	tctccccacc	cccaccccat	ctctctcccc
	acaccttccc	cccttttctc	tttgtctctc	tcttctccct	ctttcttctc	cacccccatc
	tctctctctc	ttctcttccc	acacctctcc	catctccctc	atctctttgt	ctgtctctct
	tctccctcct	tcttttccac	ccccatctct	ctgtctctct	ctctcccat	accctttccc
	tcttctcat	ctctctttgt	ctctctctcc	ttccctctt	tcttctccac	ctccaactct
30	ctctgtctct	ccacacccat	cctccttgct	cacatctgca	ccttcagctg	tcaggggatg
	tgggatgggtg	agtgttaggg	atagaggaga	tgggagagag	atgactgtcc	tagagaatag
	gggtgtcccc	ttctcccctg	gtgagggcca	gtttcatgaa	tgtgcaagct	ctgcacggac
	acagagcccc	acactcagaa	gggtctcaaa	cttagtctaa	tgcattcctg	ctgttgtctt
	gaaatttctca	ataatttttg	aacaaaaggc	cctgcatttt	cgttttgca	caagtctctg
35	aaattatgta	actggtcttc	accctggtct	ccgagaccat	cgtgtccccc	tttctgccc
	cacagggcac	gcattccccc	cttgagatg	atgttctctc	tccactagc	ttggagcagg
	gtccttaaca	ttggaaaata	aagagtgtct	tgaatcctga	agccccaccc	cttctctgca
	attggtctca	ttggccaagg	gtcaaacacc	tgtcttcaaa	ggacctagtg	tgtccctagc
	actagctctc	ccattagtcc	ccagagacaa	tgaatctctt	ctcattggct	atgggtggaag
40	tccataatct	gcaagacaaa	gaccgataac	tgaggaatgt	atgagaatga	gttgggcttt
	gatctgaagc	caaagttaat	ctccggctct	attccctcta	gggtgactca	gggggaccgt
	tgggtgtgag	aggtagccctg	caaggtctgg	tgtcctgggg	aactttccct	tgcggccaac
	ccaatgacct	aggagtctac	actcaagtgt	gcaagttcac	caagtggata	aatgacacca
	tgaaaaagca	tcgctaacgc	cacactgagt	taattaactg	tgtgcttcca	acagaaaatg
45	cacaggagtg	aggacgcgga	tgacctatga	agtcaaat	gactttacct	ttcctcaaag
	atatatttaa	accaacctca	tgccctgttg	ataaaccaat	caaattggta	aagacctaaa
	acaaaaacaa	ataaagaaac	acaaaaccct	cagtgtctga	gaagagtcag	tgagaccagc
	actctcaaac	actggaactg	gacgttcgta	cagtctttac	ggaagacact	tggtaacagt
	acaccgagac	ccttattcac	cacctttgac	ccagtaactc	taactctagg	aagaacctac
50	tgaaacaaaa	aaaatccaaa	atgtagaaca	agacttgaat	ttacctgat	attatttatc
	acagaaatga	agtgaacca	tcaaacatgt	tccaaaagta	ccagatggct	taaataatag
	tctggcttgg	cacaacgatg	ttttttttct	ttgagacaga	gtctctgttg	cttgggctgc
	aatgcagtga	tgcaatcttg	gtcactgca	acctccgcct	cctgggttca	agtgtattct
	gtgcttcagc	ctcccaagta	cctgggacta	caggtgtgca	ccaccacacc	aggctaattt
55	tttgtgtatt	tttactagag	acagggtttc	accatgttgg	ccagcatggg	cttgaacggc
	tgacctcaga	tgatccaccc	accttggcct	cccaaagtgc	tgggattaca	ggcatgagcc
	accacggcca	gcccacaatg	atattacaaa	cctattaaaa	atgatactta	gacagaattg
	tcagtattat	tcaagaacat	ttaggctata	ggatgttaaa	tgacaaaagg	aaggacaaaa
	atatatatgt	atgtgacctt	accataaaaa	aatgaaatat	tcacagaatc	agatctgaaa



```

acacatgtcc cagactgcat actggggtcg tcatgaggtg tctccttcc tctgtgtact
tttccttgaa tgtgcacttt tataacatga aaaataaagg tggggaaaaa agtctgaaga
tctaagattg gagagaggtg acctttcagg aaggagagact agaaagaaat atgtgcttg
ttttgagccc tggtcctgcc ggccctgttc cagggcatat ttccatttcc cagatctcag
5 tttttcctgt ctgtaaaatg ggagagagag gaaaggatgg agagaggaag aaggaaggga
ggaggagga gagaacaggc caacttcac agcgtgggaa ggggtgtgaa agtgtttctg
agcatctcac gagtgacaag tgaggaggga ggctggcggg ttccagaggg attgggatga
cagtagacag gacacagggg tcccacgggg gtctgccaga agtaagcaaa cagtgccgga
ggaagatggt ggcacctgct cccaagaag ggagggaaa gaaacctcggg aagcgggtag
10 gatgaggag gagtcctctg tgactcagag cctggccaca gcccagcca tctaactca
aagatcctct gtgtggtcac acctcagacg ctgctgaccg aggagccact ccagcccagg
accagccctc ctacctgttc ttctgtttt tctcccagaa ttccctcccc accaagatcc
tccagatcct tccctcctt atctcatctc cctctgagtc tctcctaacc caggcaccac
agccctgtca tattgcagaa attctgcagc cgctaattct gatttctcca tataggaggc
15 taacacagaa aacgcaggag tccaggcccc cagcccctcc ttctcagac ccaggagtcc
agaccccccg cccaacccc tctcctccta gaccaggag ccagggtccc cagccccttc
tgtttctggg cctgtcaagt ttaagaatgt caaacatttt cgaccagtca ttcccctgaa
gttttagcaa cattttctct ctcttctgca aggcactcca acattcaatc tggaaattta
aaagtaaca aaacattgca ttgtcactaa gtcagcctgg agatccctgg ccctggccct
20 ctgctctcct atacgcaagc tacaggtaga ttggtttgca atgactgaga tggtagtaat
gttgattttt ttaagtaat tcattttct ttgggtaagc agtatagtgt ggtagttaag
ggactagctc tggatcttgg ctctctgggt tcaaatccca gttctagtcc ctacaagcta
ttttccttta agctcattac ttcccctgtc cctgttcctt catccttgaa atgggagaaa
aagcacctac tttctagggt tattacagag attcaataag ttaatatata gaaagtgtc
25 aaacattgt

```

## SEQ ID NO. 13

Hk8 amino acid

```

30 MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQ
PWQAALFQGQQLLCGGVLVGGNWLTAACHCKPKYTVRLGDHSLQNKDGPQEIPVVQ
SIPHPCYNSSDVEDHNHDLMLLQLRDQASLGSVKPISLADHCTQPGQKCTVSGWGTV
TSPRENFDTLNCAEVKIFPQKCEDAYPGQITDGMVCAAGSSKGADTCQGDSSGGPLVC
35 DGLAQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSKG

```

## SEQ ID NO. 14

KLK8 nucleic acid

40 CDS 35..817

```

gtgaccccg cctggattc tggagacct caccatggga cggcccgac ctctgctggc
caagacgtgg atgttctctg tcttgctggg gggagcctgg gcaggacact ccagggcaca
ggaggacaag gtgctggggg gtcatgagtg ccaaccccat tgcagcctt ggcaggcggc
45 cttgttccag gccagcaac tactctgtgg cgggtgcctt gtaggtggca actgggtcct
tacagctgcc cactgtaaaa aaccgaaata cacagtacgc ctgggagacc acagcctaca
gaataaagat gggccagagc aagaaatacc tgtggttcag tccatccac acccctgcta
caacagcagc gatgtggagg accacaacca tgatctgatg cttcttcaac tgcgtgacca
ggcatccctg ggttccaaag tgaagcccat cagcctggca gatcattgca ccagcctgg
50 ccagaagtgc accgtctcag gctggggcac tgtcaccagt ccccgagaga atttctga
cactctcaac tgtgcagaag taaaaatctt tcccagaag aagtgtgagg atgcttacc
ggggcagatc acagatggca tggctctgtc aggcagcagc aaaggggctg acacgtgcca
gggcgattct ggaggcccc tgggtgtgtg tgggtgactc cagggcatca catcctgggg
ctcagacccc tgtgggagggt ccgacaaacc tggcgtctat accaactctt gccgtacct
55 ggactggatc aagaagatca taggcagcaa gggctgattc taggataagc actagatctc
ccttaataaa ctcacaactc tctggttc

```

## SEQ ID NO. 15

KLK8 nucleic acid

5 CDS join (&lt;1..39, 418..712, 878..&gt;946)

Exon &lt;1..39

Exon 418..712

Exon 878..946

```
10 tcttcggttc ccggttactg gcagcagccc cctcctccca caaaagatca ggttccaagc
   ttctcctttt aaaagtactt agaattcagc cccagctctt ctctccctc acaccagga
   atccaggccc ctagccctc ctccctcaga cccaggagtc ctggcccta gcagccctt
   cctccctcag acccaggagt ctgggcccc agccctcct cggtcagacc taaatccag
   gtccagttcc ctctccctt agatttagga gtccaggccc ccagcctctc ctccctcaga
15 cccagggaatc caggccccca gcctcctccc ctctcagaac taaaatcttg gccccagcc
   ctttatgttt cagatcgtag agtctcagca ccgagtcctt cctctcccta gcctcaggag
   tcttagattc cagccctccc tccctcaaga ttccacgttc aatccctcc gcccttctc
   actcacacc agtggtccag ttcccagaag ctcccaggc tctagtgcag gaggagaagg
   aggaggagca ggaggtggag attccagttt aaaaggctcc agaatcgtgt accaggcaga
20 gaactgaagt actggggcct cctccactgg gtccgaatca gtaggtgacc ccgcccctgg
   attctggaag gtgaggtgca gaggtactca gatacagaca tcaggccccg gaccctcctt
   ctccagattc caggacccca gcctcagatg cccttctctg tcgagatcca gcagtctgga
   ccccggttcc ctctcctccc taatttagga gtcccagctc ccagctccct gtccctcag
   acccagacat cgaggactcc cccctccctt ggaatgtagg aatccagttc cccagcctcc
25 tccttctccc agagaagccc agaacagccc cagatactct cggctgcctc cccagtgcc
   aaatccagaa ctgggagctc aggtcctccc ttctgttta ccggccccgc cctctccatt
   tcccagacct caccatggga cgccccgac ctctgtcggc caagacgttg atgttctgc
   tcttgctggg gggagcctgg gcaggtgagg aggggttgcg aggcctccg aggggagga
   tctgaaggca gcagtggcg tggtgagttc gtgggaatgc cgcgggggtt atgtgggtgc
30 gtgtgcacgg atgtgaagag tgcgatacgg tgcaggagcc tctgtgggtt ttctcaggg
   tggacagagg caagaaacag gtagcagcag gtaggagtag gttccgtgat gctgtaaatt
   gtctgaatag ctacagcctt tgggggctgc ttgcttgggg gcatagattc acctgggagt
   actcggggcc tgtagactca tgtggaagca tgtgggggca ttcttgggtg tgtgactctt
   gtatgatgac acatggactg aaatgagttt ccccggtgtg cagcgtgtgg aagcctggac
35 ctctcacta agttgtatgc ggagaacttg ccgtgtgtcc atttgaacc acagtggcct
   tcccagccct cgactgccc cagagggttg cgatccaacc ctctccctcc tgcgtcagga
   cactccaggg cacaggagga caaggtgctg gggggtcatg agtggcaacc ccattcgcag
   ccttggcagg cgcccttgtt ccagggccag caactactct gtggcggtgt cctgttaggt
   ggcaactggg tccttacagc tgcccactgt aaaaaaccgt gagtggatga tgggggcaga
40 ggtcagctgg ggcttaagga aagagggggc tgggggttct actcaggaag gagagagctg
   aggactggac tcctgggtct gaaggaggag ggggctgggg gcaatacccc tgccctgggtc
   ccaaactatc cccaccatta caggaaatac acagtacgcc tgggagacca cagcctacag
   aataaagatg gccagagca agaaatacct gtggttcagt ccatcccaca cccctgctac
   aacagcagcg atgtggagga ccacaacat gatctgatgc ttcttcaact gcgtgaccag
45 gcatccctgg ggtccaaagt gaagcccatc agcctggcag atcattgcac ccagcctggc
   cagaagtgca ccgtctcagg ctggggcact gtcaccagtc cccgaggtag tgggcttgtc
   cactaatggg agggagagga ggagctggtt ggcccagtg aaccaagct attggcaag
   cttggtcccc cagagggaga caaagaagg aaagtgatca tgatgttgag attcacaagc
   aggagtcata tgagaagcct cgaagatctg actactaaca agagtgggtg agaaagaac
50 caactagcag attgttaagc agggccagaa aagcccatcc tgtatggcgg agagcactac
   octatgggac ctttggttgt ataggatttt atataatctt aacacacttg ggatattttg
   gacttctcag agggccagga aaacaggctc ctaagcacct tctccccac ctccctcctt
   tttttttttt tttttttttt ttttgcagg ggacacagtt tcaactcaat gcccaggctg
   gagtgcagtg gcacaatctc agctcactac aatctctgcc tcctgggttt aagcgattct
55 cgaacctcag cctcccgagt agctgggatt acaggcacc gccaccacgc gagctaattt
   ttgtatttgt ttagagaca gggtttcgcc atgttgacca ggctggtct gaactcctga
   cctcaagtga tctgcccgc tcggcctacc aaagtgtctg gattacaagc atgagccacc
   gccctgggc tttttctcct tcttgaacct aggaaacctg ggccctggtc acctcttccc
```

tcagacccgg gagtctaggg ccttcttctc ccaggaccca gcagtcctag ctccctcttg  
actctggacc ccaaaatctg gacctccaat gaagctgtcc ctttgggact ccagaatcca  
gagagccctt cgccttctt cactgtgaaa acattgggac tcacctaa gaagcaagga  
gcttctccag gaagtggcaa agtcagcatt caggctccctg cctgcctcac tctgtctctg  
5 aatgcttttg atgagacagt ttgctggctgt ggaaacacac gtgctcgcaa atcaagtaga  
tcagttcaaa ctatggctct gccctttctc cctgggcaaa ttcttttcca tctctgagcc  
tcactttcct tatctgcaaa atgggaatca ttaaccaata gatttttttag ctacgtgagg  
gttcagggtta catcagtttt cttcattgtg ttctcctagt gctcagaatg gtgcctgcta  
catggtagat gctcaataaa tatttgttga atggatgacc tgatgaataa ataataaat  
10 gaatgaatgc atctctcctt caaagcgtg ttgtgaggat taaatgagat tatgagcata  
ttgcttttgg cacgtagaga tgccagatgg aagtttttcc ccctcaaaga ggctttggag  
aagtctattc ctcaaaagag gttaataaaa aagatcaatt ccaccttcaa tcatataattc  
aactcttatt tactgagcac ctagtatgcc tgaggtgctg ttgcaggcgc tggatataca  
gcatgagca aactgtacaa agtccttctc gttatggagt tgcaagctag gtgggagaaa  
15 tagacaataa acaatacac ataaaaataa acattaggca aagtgctgta aggaatatt  
gtggagcgta aaaggatagg gagtaatgga gggatggat tttttaattc gagggtcag  
ggaaggcttc caggaggagg tgacacttgg aaggagcagg atttagcaca gattgaggac  
ccggttgtct gagggcagga agactggaaa gagagggtca gaggatcat aaaggggcac  
ctccaagtaa ccccagccc cttgatttga aaatggctca gggtaagaaa aaacacgtga  
20 gatccaaggg cccctctcta gatggagaaa gcccaatagc aacaagtaca gcttcgttta  
atgtggtgag aagtgtgtc cctgtgcac agtgcagaa aattcccca tgcagctgga  
aaactcccc attacatcct ggaaagaaag ggggttagat ccagacaggg atggaggcaa  
agggtgtct tctcaggga ccttacaacc tctccccct cagagaattt tctgacact  
ctcaactgtg cagaagtaaa aatctttccc cagaagaagt gtgaggatgc ttaccgggg  
25 cagatcacag atggcatgg tctgtcaggc agcagcaaaag gggctgacac gtgccagggtg  
agcaatttct gaaatccttc tctcacaca tccctcattg ccctctcgag gttcaaggct  
tggtatgggg tggggtgggt aaggagtgga ccccaaagac ttggcaccog gagtgttcac  
ccctatctct acggattggg agccaggttc agagaagcca aactctctct ctgaaagtca  
cactgcatag ggattaggag cattgaattg ctgctgtgt tttctttcga acgcttaact  
30 acaggatggg atgcagagtt gggggctata gaggggtggg tagatgtcca gcaaggagag  
agtctgaggg tagagtttag acgagttgcc tgctctctgg ttccagcat tacatcgtgg  
aatttgtgcc gactttacca ccggcccggt ggtcagccac tttcttttga gaccgggcta  
gaatcccaag gctggttccc ttctcctgtg attggttcc tgggagacaa tgggtgtctc  
ccagttgggt ggagtaaatg gtgccattga cttcagtgct tctcaatga gctccaatcc  
35 cttctctctg ggatttgtca tagatataac tctctcttct aactgcaact tctgtgttcc  
tagcactttc ccttggtctt gtccacttcc tggaaagccc accaccttg ccaatgactg  
gtccctaaat acaatgcttt ctccccaatt ggccaaaaat ggagtcgttt ccacagcga  
tactgccatg aaagccagtc tctggattgt tctgtagaga tagtggtctc ttccacaaat  
atctcagcca tggctctctg gggataggt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt  
40 gtgtgtgtgt gtgtccataa gaatcttgat cctttctcct atgttttgca actgcaatca  
atgatgcctt cactattggc caggaacaga ggaaacttca gctcagtcct ctccaagtaa  
tccctactgt cttctccctg gattggaccc tcgagaactc tttttttttt ttttttgaga  
cagggtctcg ctctgtcccc caagctggaa tgcagtgga caatcttggc tccactgcagc  
ctctgctccc cagttcaagc aattctccca cctcagcctc ccgagtagct gtgattacag  
45 gtgtgcacca ccaccccag ctagtttttg tatttttagt agagacagg attaccatg  
ttggccaggc tggctcggaa cgcctgactt caagtgatct accgcctcgg cctcccaaag  
tgccgggatt gcagggtgta gccaccaagc ctgtctggga ttctattctt tccctctctc  
tgtcagtggt ttgaccacta cccttagaca ccattgtctgt ctgtacatgg aagcccaaag  
ccctgtcctg actggtctca ggggacaatg cttttacccc cattggctac aggggaccaa  
50 tcatgccaaa gaactggtaa aacgtggga cagcaggaaa agggacgttg tggacatctc  
agatgcaagg ctgttctcat tctccctgtc tagggcgatt ctggaggccc cctggtgtgt  
gatggtgcac tccagggcac cactcctctg ggctcagacc cctgtgggag gtccgacaaa  
cctggcgtct ataccaacat ctgccgtac ctggactgga tcaagaagat cataggcagc  
aagggtgat tctaggataa gcaactagat tcccttaata aactcacaac tctctggttc  
55 cttgcctgtc tctgttttgg cctgtgggg agggatccgg ggggatccgg gatgttctc  
gctggcagac taggtgggga tgtgcagaaa ccaagttctc atggtaactt tggccatcac  
cactgcctaa agtgatccct cgttttctgg aagaacttgg gtaagagctt tatttcaggg  
gagaaaatac atacaaaggt cttcaaacat tcagtgggct ggtatgtgaa agacagttt  
gaaagagttt gtgttttagt ttctgtagca aagcatttac aagcttttga gataaaattt

tcccccttta aaaaaataatt

SEQ ID NO. 16

5 Hk10 amino acid

MRAPHLHLASAASGARALAKLLPLLMAQLWAAEAALLPQNDTRL  
PEAYGAPCARGSQPWQVSLFNGLSFHCAGVLVDQSWVLTAAHCGNKPLWARVGDDHLL  
LLQGEQLRRTRRSVVHPKYHQSGPILPRRTDEHDLMLLKLARPVVPGRVRALQLPY  
10 RCAQPGDQCQVAGWGTTAARRVKYNKGLTCSITILSPKECEVFYPGVVTTNNMICAGL  
DRGQDPCQSDSGGPLVCDLTLQGILSWGVYPCGSAQHPAVYTQICKYMSWINKVIRSN

SEQ ID NO. 17

15 KLK10 nucleic acid

Gene 1...1580  
CDS 220...1050

20 catcctgccca cccctagcct tgctggggac gtgaaccctc tccccgcgcc tgggaagcct  
tcttggcacc gggaccgga gaatccccac ggaagccagt tccaaaagg atgaaaagg  
ggcgtttcgg gcaactggag aagcctgtat tccagggccc ctcccagagc aggaatctgg  
gaccaggag tgccagcctc acccacgcag atcctggcca tgagagctcc gcacctccac  
ctctccgccg cctctggcgc ccgggtcttg gcgaagctgc tgccgctgct gatggcgcaa  
25 ctctggggccg cagaggcggc gctgctcccc caaaacgaca cgcgcttga cccgaagcc  
tatggctccc cgtgcgcgcg cggctcgcag ccctggcagg tctcgtctt caacggcctc  
tcgttccact gcgcgggtgt cctggtggac cagagttggg tgctgacgc cgcgcactgc  
ggaaacaagc cactgtgggc tcgagtaggg gatgaccacc tgctgcttct tcagggagag  
cagctccgcc ggaccactcg ctctgttgtc catcccaagt accaccagg ctcaggcccc  
30 atcctgcaa ggcaacgga tgagcacgat ctcatgttg tgaagctggc caggcccgt  
gtgctggggc ccgcgctccg ggccctgcag ctccctacc gctgtgctca gcccgagac  
cagtgccagg ttgctggctg gggcaccacg gccgcccga gaggtaagta caacaagggc  
ctgacctgct ccagcatcac tatcctgagc cctaaagagt gtgaggtct ctacctggc  
gtggtcacca acaacatgat atgtgctgga ctggaccgg gccaggacc ttgccagagt  
35 gactctggag gccccttgt ctgtgacgag accctccaag gcatcctctc gtgggtgtt  
tacctctgtg gctctgccc gcatccagct gtctacacc agatctgcaa atacatgtc  
tggatcaata aagtcatacg ctccaactga tccagatgct acgctccagc tgatccagat  
gttatgctcc tgctgatcca gatgccaga ggctccatcg tccatcctct tcctcccag  
tcggctgaac tctccccttg tctgactgt tcaaacctct gccgccctcc acacctctaa  
40 acatctcccc tctcacctca ttccccacc tatccccatt ctctgcctgt actgaagctg  
aatgcagga agtggtggca aaggtttatt ccagagaagc caggaagccg gtcatoaccc  
agcctctgag agcagttact ggggtcacc aaactgact cctctgccac tccctgctgt  
gtgactttgg gcaagccaag tgccctctct gaacctcagt ttctcatct gcaaatggg  
aacaatgacg tgctacctc ttagacatgt tgtgaggaga ctatgatata acatgtgtat  
45 gtaaactctt atggtgattg tcatgtaagg cttaacacag tgggtggtga gttctgacta  
aaggttacct gttgtcgtga

SEQ ID NO. 18

KLK10 nucleic acid

50

Gene 1..5574

mRNA join(48..120,605..701,2455..2635,3589..3863,4195..4328, 4793..5474)

CDS join(614..701,2455..2635,3589..3863,4195..4328,4793..4945)

Promoter 1...47

55 5'UTR join(48..120,605..613)

exon 48...120

exon 605...701

	ttgggggtcaa	aaggggaaggt	cccgccagg	gtccctgggc	agaggatacc	agcggcagac
	cacaggcagg	gcagaggcac	gtctgggtcc	cctccctcct	tcctatcggc	gactcccagg
5	tgaagctacc	tgacccccac	ccgggttggg	gtggattgcg	agaggatggg	tgggaacccc
	cgggccacag	gcaggagccg	gcttagagcc	tcggtttctc	cactgcggga	cgcggaagtc
	ccccgcgtgt	gaggttgaga	aagaggctcc	cactggctcc	gagcctcggg	tccccacccc
	gcgcgtgaag	gaggggaaa	cctcgggcgc	ggcgtggctg	cagccagggt	agctggggcg
	cgagagcgc	tccactcggg	cacagggagg	acgggaagat	gccgcgagg	gcgtcattag
10	ggtaattgtg	cccattaccg	tggtccagcc	cgaatctccc	gtctgccagc	ccctgggggt
	atcggtgca	ggtcaaaagg	gtgcttgccg	gctgggggtg	cccaggctgg	gcggatgagc
	gcggcgagg	gggggtctct	aacggagcat	ctgttttaac	ccgccccctg	acacacccca
	gcagatcctg	gcatgagag	ctccgcacct	ccacctctcc	gccgcctctg	gcgcccgggc
	tctggcgaa	ctgctgccgc	tgctgatggc	gcaactctgg	ggtaaggtag	gggacagggg
15	gcggggagag	gcgccggtgg	gaggcacggg	cgaggaggca	atgtgttccc	gtcaccaagc
	cccgcgacc	tctcctcccc	cgagacccca	gcacccaccc	agcgtccgg	agcacggccc
	gcccccaagt	cagctggggc	cttctctctg	ctcgccccct	gggtgaccgc	ccccactcag
	gcctgtccg	atttctgcca	cccgatccct	cgctcttccg	tggactcttc	ggcgtgttct
	ttctcctcct	cttacgcgcg	cccatcccg	cccgctcca	ttttacacta	ccgtgttttg
20	ttttgtttgt	ttgtttgttt	gagacggagt	ctcgtctgt	cgccaggct	ggagtgcagt
	ggcgcgatct	cggctcactg	caagctccgc	ctcccggtt	cacaccattc	tcctgcctca
	gcctcccag	tagctgggac	tacaggcgcc	cgccaccaag	cccgctagt	tttttgtatt
	tttagtagag	atgggatttc	accgtggtct	cgatctcctg	gcctcgtgat	ccgcctgcct
	cggcctccca	aagtgcctgg	attacaagcg	tgagccaccg	ctcccgcca	caccacagt
25	ttttatcctg	agtcttgct	taaccgtttt	tgccctctcc	cctcactttt	tttcttctc
	tttccctttc	tctctotttt	ttotttttct	ttotttcttt	tctttctttc	gttcttcttt
	tccttccttc	ctttccttcc	ttcctttcgt	ttgtttcttc	tttctgtttt	tactttctct
	cttatttttt	cttcttttct	tcttgttttt	tttctttctc	tctctttttc	tttcttctt
	tttttttttt	tttttttgag	acagggcagt	gctctgtctc	cgtggctgga	gtacagtggc
30	ccaatcagag	ctcactgcag	cctcgacctc	ctgggctcaa	gcgatactca	gcctccagag
	tagctggtac	cacaggcatg	caccaccaca	tccggctttt	tttttttttt	tttttttttt
	tttttttgag	acagggtctc	actctgtcgc	ccagactgga	gtgcagtggc	ccaatctcgg
	ctcattgcaa	cctccacctc	ctgggctcaa	gcgatcctcc	cacctcagcc	tcccaggtag
	ctgtgactac	aggcgcatgc	ctccgcgact	acttttttgt	tggtgtgtgt	gttgtttgt
35	ttttgtagag	actgggtctt	gctgtgttgc	ccgggctggt	cttgaattcc	cgagctcaag
	cggctccaacc	gcctcggctt	cccaaagtgc	tgggattaca	ggcgtaaacc	actgcgcccc
	acccctctcc	tggttttcaa	tcccgttttg	ttattcacac	cccttctctt	ccccgatccc
	cgagttctat	ccccgcaccc	ttacctcccc	gccgcgttca	atccccgccc	ctctatcgac
	cagcgacgtt	ctagccagct	ctccaggcgc	gctgcgttca	gtccctgccc	tccagaccca
40	ccctattctg	tctcattact	ccacgctacc	ctatcccagc	ttccttccac	tttcacgcgc
	ttcttctcct	cccattcctt	cggtgcacgc	gaaaccccca	atatttccct	accaccctcg
	cgttctggct	gcgtccccc	tccccgaacg	cagtccagt	ccacagccca	gtcccaaccc
	caaaccocaga	ccctgcccctc	cgtgcttga	ttccgtcccc	tttctttctc	ccagccgcag
	aggcgccgct	gctcccccaa	aacgacacgc	gcttggaacc	cgaagcctat	ggcgccccgt
45	gcgcgcgcgg	ctcgcagccc	tggcaggtct	cgctcttcaa	cggcctctcg	ttccactgcg
	cggtgtcct	ggtggaccag	agttgggtgc	tgacggccgc	gcactgcgga	aacaagttag
	aggagatcca	tccccgagga	cgccacgggg	ggctgtggag	gcggcctcca	gggaggagcg
	ggcaggggcg	ggtctctcgg	aactcccaca	gctggagggt	cggtcccccg	tgtcttcca
	gcaggaggag	agaccgggct	tgcgctgtgg	ccgcccagg	acggtgtggt	ccttttccgc
50	atttctgggc	ccgtcaactcc	tctcccgctg	attctccttg	agctctagga	ggaggtgggt
	gcctgcaacg	gatagaagcc	agggtccggt	gtggtcagga	atttagaact	aaaggaaaag
	gttcaacttcg	tgagtccccg	ttgaaggagg	aagggttggg	tattaccaca	gagaaaatgt
	ggagttgggc	tgggctcgg	ggctcacgcc	tgtaatccca	gcactttggg	aggccaaggc
	gggtggatca	cctgaggtca	ggagttcaag	accagcctgg	ccaacatggt	gaaaccactt
55	ctcaactaaa	aataaaataa	aataaaaaat	aaaaaatttt	aaaaaaatta	aaaaaaaaaa
	aaggctgggt	gtggtggcgg	gtgcctgtaa	tcccagctac	tcaggaaagg	ggctgaggca
	ggagaactgc	ttgaacctgg	gaggtggagg	ttgcagttag	ccgagattgc	gccattgcat
	tctagcctgg	gtgacaagag	cgaaactctg	tttcaacaaa	gaaaagaaaa	agagaaaaga
	aatgtggaa	ggcttaccta	ggtgtccagg	ccccagccc	tcctcaattc	ctgcagatcc

tcagagctca aacaactgat tctctctccc catgtccact gaggtcccoct tctcccacaa  
 ggccctcttc cctcagactc ttcctatctc caggccctgc ttcactgccc acctgctttc  
 ccagtccttg tgaagggttt gccttcacat gcctcttccct tccccaggc cactgtgggc  
 tcgagtaggg gatgaccacc tgctgtctct tcagggcgag cagctccgcc ggacgactcg  
 5 ctctgttgte catcccaagt accaccaggg ctccaggcccc atcctgcca ggcgaacgga  
 tgagcacgat ctcatgttgc taaagctggc caggcccgta gtgccggggc cccgcgtccg  
 ggccctgcag cttccctacc gctgtgctca gcccgagagc cagtgcaggg ttgctggctg  
 gggcaccacg gccgcccga gaggcaagag ctggggctct gaggccagaa cctcaggagg  
 agggggctga gggcctgaac cctgggtct gaggaaggat gggctgggga ctggattcct  
 10 ggatctgagg gaggacgggc tggggtccta gatgcctggg tctgtgagtc tgaggggagg  
 aggggctggg ggcttgact cctgggtcta agtggggagg ggctggggcc aggattcttg  
 agtctgaagg aggaggggct ggggcttagg atagaaacgg tcttgtatct ggactccttg  
 ctccccagg attgggggct ggaccaggg attactggca tattctccct tcagtgaagt  
 acaacaagg cctgacctgc tccagcatca ctatcctgag ccctaaagag tgtgaggtct  
 15 tctaccctgg cgtggtcacc aacaacatga tatgtgctgg actggaccgg gccagggacc  
 cttgccaggt aggtctgaa caggagagat ctctgactcc tgggaggag gacaggagg  
 ttatgggaaa agagcagacc ctgtgccga tcccaaactc cattcccaa cccatccttg  
 accccaactc ttaccagac ctaacccct cctcatccct atcctcaatc ccatttccat  
 cctaacccca cccattccc atctccaagc ccatttccat cccctcacct tccatgaact  
 20 acaatcccaa cccaagtctc actgtgcctt cattctcatc cccagccca acctccata  
 acctgaagtc cacctccatt cctaccttc agtccatacc taattccaac cccatcccat  
 cctcgtcttt atcccaaccc aaccccttc ttccccacca ctgcccaga tccaaagtg  
 acagctctca cgttggcaca tttattgat ctctcctttc tgccaccccc agagtactc  
 tggaggcccc ctggtctgtg acgagaccct ccaaggcatc ctctcgtggg tgtttacc  
 25 ctgtggctct gccagcatc cagctgtcta caccagatc tgcaaatata tgcctggat  
 caataaagtc atacgctcca actgatccag atgctacgct ccagctgac cagatgttat  
 gctcctgctg atccagatgc ccagaggctc catcgtccat cctcttctc cccagtcggc  
 tgaactctcc ccttgtctgc actgttcaaa cctctgccgc cctccacacc tctaaacatc  
 tcccctctca cctcattccc ccacctatcc ccattctctg cctgtactga agctgaaatg  
 30 caggaaagtgg tggcaagggt ttattccaga gaagccagga agccggtcat caccagcct  
 ctgagagcag ttactgggtt caccacaact gacttctct gccactcccc gctgtgtgac  
 tttgggcaag ccaagtgcct tctctgaacc tcagtttct catctgcaa atgggaacaa  
 tgacgtgctt acctcttaga catgttgtga ggagactatg atataacatg tgtatgtaaa  
 tcttcatgtg attgtcatgt aaggcttaac acagtgggtg gtgagttctg actaaagggt  
 35 acctgttctc gtgatctgac cagtcctcg tgaaagcgtg tgtccaggga agaagtgcac  
 agggtagccc ccaatcccaa ccttccatcc ccaaccctta gggatgatgg aaga

## SEQ ID NO. 19

40 Hk11 amino acid

MQRRLRLRDWKSSGRGLTAAKEPGARSSPLQAMRILQLILLALA  
 TGLVGGETRIKGFCEKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAHCLKPRYIV  
 HLGQHNLOKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITWAVRP  
 45 LTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRCANITIIHQKCNAYPGNITDTM  
 VCASVQEGGKDSQGDSSGGLVLCNQSLQGIISWGDPCAITRKPVGVTYKVCKYVDWIQ ETMKNN

## SEQ ID NO. 20

50 Hk11 amino acid

MRILQLILLALATGLVGGETRIKGFCEKPHSQPWQAALFEKTR  
 LLCGATLIAPRWLLTAAHCLKPRYIVHLGQHNLOKEEGCEQTRTATESFPHPGFNNSL  
 PNKDHRNDIMLVKMASPVSITWAVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHT  
 LRCANITIIHQKCNAYPGNITDTMVCASVQEGGKDSQGDSSGGLVLCNQSLQGIIS  
 55 WGDPCAITRKPVGVTYKVCKYVDWIQETMKNN

## SEQ ID NO. 21

## KLK11 nucleic acid

```

aggaatctgc gctcgggttc cgcagatgca gaggttgagg tggctgcggg actggaagtc 61
atcgggcaga ggtctcacag cagccaagga acctggggcc cgctcctccc cctccaggc 121
5 catgaggatt ctgcagttaa toctgcttgc tctggcaaca gggcttgtag ggggagagac 181
caggatcatc aaggggttcg agtgcaagcc tcaactcccag ccctggcagg cagccctgtt 241
cgagaagacg cggctactct gtggggcgac gctcatcgcc ccagatggc tcctgacagc 301
agcccactgc ctcaagcccc gctacatagt tcacctgggg cagcacaacc tccagaagga 361
ggagggctgt gagcagaccc ggacagccac tgagtccttc cccacccccg gcttcaacaa 421
10 cagcctcccc aacaaagacc accgcaatga catcatgctg gtgaagatgg catcgccagt 481
ctccatcacc tgggtgtgtc gacccctcac cctctcctca cgctgtgtca ctgctggcac 541
cagctgcctc atttccggct ggggcagcac gtccagcccc cagttacgcc tgcctcacac 601
cttgcgatgc gccaacatca ccatcattga gcaccagaag tgtgagaacg cctaccccgg 661
caacatcaca gacaccatgg tgtgtgccag cgtgcaggaa gggggcaagg actcctgcc 721
15 ggggtgactcc gggggccctc tggctctgtaa ccagtctctt caaggcatta tctcctgggg 781
ccaggatccg tgtcgcatca cccgaaagcc tgggtgtctac acgaaagtct gcaaatatgt 841
ggactggatc caggagacga tgaagaacaa ttagactgga cccaccacc acagcccatc 901
accctccatt tccacttggg gtttggttcc tgttcaactc gtttaataaga aaccctaagc 961
caagaccctc tgcgaacatt ctttgggcct cctggactac aggagatgct gtcacttaat 1021
20 aatcaacctg ggggtcgaaa tcagttagac ctggattcaa attctgcctt gaaatatgtt 1081
gactctggga atgacaacac ctggtttgtt ctctgttgta tccccagccc caaagacagc 1141
tcctggccat atatcaaggt ttcaataaat atttgctaaa tgagtg

```

## SEQ ID NO. 22

## 25 KLK11 nucleic acid

gene 2313..7622

30 mRNA join(2313..2398,4189..4263,5061..5217,5545..5810,  
6627..6763,7158..7622)

CDS join(4224..4263,5061..5217,5545..5810,6627..6763, 7158..7310)

```

35 tgataaatagt gttctctctc ctcatgtgtc agggccccag ccattgtcct tgagagaatg 61
ctcgactctt tatgttgtct tgacagcctc ccttgagatt ggtcattaat gactgtgtc 121
tctctcctca ttggtcaggg ccccagccat tgccttgag agaacctctg tcctttatgg 181
agttccaccc ttcttccctg ggattggccc cttagagacag tggttcttct cttttgggta 241
gccattgcca ttgtcctccg ggaaagtgat tatactcttt tgtctaataga ccagacttgg 301
40 agccctcccc aaggcccagg actgggttga agggttgggg aggaaaacag aaataagatg 361
tctcccttgt tcagacagta cttctcttcc cttccagggt gattctgggg gccccctggt 421
gtgtggggga gtccttcaag gtctgtgtgc ctgggggtct gtggggccct gtggacaaga 481
tggcatccct ggagtctaca cctatatttg caagtatgtg gactggatoc ggtgatcat 541
gaggaacaac tgacctgttt cctccacctc cacccccacc ccttaacttg ggtaccctc 601
45 tggccctcag agcaccaata tctcctccat cacttcccct agctccactc ttgttggcct 661
gggaacttct tggaaacttta actcctgcc a gcccttctaa gaccacagag cggggtgaga 721
gaagtgtgca atagtctgga ataaatataa atgaaggagg ggccatgtct gtccatttga 781
agtcctcatg ctggttgaga ctggaagaag gactcagcag ttccctatc tcataggagt 841
agaaacagag ctcaaataag gccaggcaca gtggctcaca cctgtaatcc catcactttg 901
50 ggaagctgag gcaggtggat cacctgaggt caggaaactc ggaccagcct ggtcaacata 961
gtgaaacccc aactctacta aaaatgcaaa aattagccag gcatggtggc gcatgcctgt 1021
aatcccagct actcaggagg ctgagacagg agaatagcat gaaccctgta ggcagaggct 1081
gcagcgagcc gagattgaac cattacactc cagcctgggc gacagagcga gactccatct 1141
caaaaacaaa caaacaaaaa acccagtgtc caaataggat gagggcttcc cctgagtagt 1201
55 tactcagaaa tggagtagaa aaagttactt ttaataatat aggccgggtg cagtggccca 1261
cgcctgtaac cccagcactt tgggaggccg aggtgggagg atggcttgag ctcagatttc 1321
gagatcagcc tggcaacaca gtgaaatctt gtactacaa aaacacaaaa aattagctgy 1381

```

	gtgtgggtgt	gcgtgcctgt	agtccagct	acttgggaag	ctgaggtggg	aggatcacc	1441
	gagccgggga	ggtggaggct	gcaaagagcc	gagatcatgc	cactgcactc	cagcctgggc	1501
	aataaagtga	gacctgtgtc	caaaaacaaa	aaccagcaa	tataaataag	acacatgttt	1561
	cttcatctgg	cataatagaa	atagtgccca	gagcttataa	gcttttcaag	agtccacaaa	1621
5	agacccgaaa	aagaaaaaga	aaattgttag	ctccaaaata	ccagatgaaa	gctgcaaagt	1681
	caacattttat	gaccatttta	tcacatgtcc	ataaaacgta	gcattctttc	cactagccaa	1741
	ctgcagttta	ctttcttgta	atgaagcata	cattgtatct	ttaatgtggg	acgtggcttt	1801
	gttctaataa	gacgaagggt	ggagtgcagg	cttggaaagc	aggagagctc	agcctacgtc	1861
	tttaatcctc	ctgcccaccc	cttggattct	gtctccactg	ggactcaaga	ggtgaggaga	1921
10	gaccatctcc	ccaaatgcac	tgaagggaaa	ctggaggagg	gagggagtga	ggggtgatca	1981
	taccagcgga	ggcacatttg	ctgagccccc	ccgcagtctg	ctctttccaa	gtggaccctc	2041
	ctggaagcct	gatcccaacc	tcccctgcaa	gcaggtctgt	cacccccatc	tctcagatga	2101
	agaaactgag	ccttgcaggg	gtggagtccc	ttgtccccc	gtcataaggg	tagtcatagt	2161
	agtaggaaga	ggaagcacct	aggtttgagg	ccagggtggt	ctgctgtcag	aacctaggcc	2221
15	ctcccctgcc	ttgtctccca	cctgggtcagg	ggagagaggg	gaggaaagcc	aagggaaggg	2281
	acctaactga	aaacaacaa	gctgggagaa	gcaggaatct	gcgctcgggt	tccgcagatg	2341
	cagaggttga	ggtggctgcg	ggactggaag	tcatcgggca	gaggtctcac	agcagccagt	2401
	aagtgaacag	ctggactcgg	gctgcctggg	cggcagggag	aagcggggcag	gggaagggtc	2461
	agcagaggag	cgaggcccca	gaggagccct	ggggtggagc	acagccaaag	gctctgtttc	2521
20	ctttcctgga	ctcggcttcc	acaggccctg	acctgcctcc	cccaccctcc	ggctctgccc	2581
	ctgtgcctgg	cagcagcccc	acctgtgtga	catcccagca	cacccccctc	ctccttgcaa	2641
	aggagaaggg	agcggcctag	gggaggccag	gggcccacct	gggctggggc	tgtggagagg	2701
	gagtggctgg	gacgggagga	aaaagagaga	cggagattag	atggaagaag	agggatttca	2761
	agacaaattg	ccagagatgc	agtcagagag	actgactgag	agacacaaa	atagaaggaa	2821
25	ttagagaaag	ggccacacag	agccagacag	agagagaaga	gtggagatgg	agacagggag	2881
	gaggacagag	aaaggcagac	agacacatag	ggacagaaag	agaaaaatca	cacaaagtca	2941
	gaattactga	atgacaggga	atgacacata	gaacgagaca	cagattcaga	gactcagggc	3001
	agggaaagga	aggctgcaga	cagacagaca	gacagaggga	ggctgagaca	cagggagaa	3061
	aggggcttgg	agaggtggca	caggcaggca	gccagtgcct	cagaggcctc	cggggagggc	3121
30	cctcacacac	accccgcccc	ggggcattaa	ggcagggcct	ggaggccaag	catcctgggc	3181
	ccgcccaggg	ccgccccctc	gccagcccgc	ctgcctgggt	cctggcacct	ggcgctccaa	3241
	cccagcctac	ctgctgtagc	tgccgccact	gccgtctccg	ccgccactgg	gccccagag	3301
	ccccagcccc	agagcctgtg	agtccaggag	gaaagggaa	ctgccccctc	ccgtccagg	3361
	gtcagccctc	cccaaggaca	cctgtcccac	tcgggcaccc	atttctccct	ctgctctgtc	3421
35	cttctctgct	tgggtggggg	ttcctggcct	ctctctacac	ctctcacctc	cgatggctgt	3481
	ccgcagcctc	agttacctct	aatctccatg	gcttcagctg	ctgagctggc	cctctgctcc	3541
	cacccccgct	ggccagggca	gcggagggca	ctggccctcc	cctcgaccag	cccogcccag	3601
	ctttgcttgg	ctgtccttca	aaagggcagg	ggtttggcgg	acagggcttc	agcaagccgg	3661
	gtgatggggg	tcccagacat	tgtctggggc	tgagccccct	actccccctc	agcagacctc	3721
40	aaaggctcca	tatcgctctg	ctgcgaagac	aatgaaaaag	gggtgggtac	ggaacgggtg	3781
	ctggttcccc	ttgtccttcc	accccaagct	gctggggcct	ggccagctct	caaggcaaga	3841
	aggaaaacat	cctctgacat	gtgcggggga	ggtcccatgg	ctgacttgaa	caggggccga	3901
	ccatggcttg	acagctcaaa	gccccctcca	acgacttcca	catggttctt	ggtatctcgg	3961
	aagcttctag	ctgtgaccag	gccctctcca	aggccaccct	agacacctaa	gatataat	4021
45	aagtgttttg	agatctgagt	gctgtgagaa	acaggggatt	tcccaacct	tgtttctccc	4081
	aagtggggag	cgaggagcag	tgagggagag	aggagagggc	atgagccagc	ccccccctcc	4141
	cgatttcccc	gtaaagtgat	gcggccccat	gtccctcctt	gttcccagag	gaacctgggg	4201
	cccgtctctc	ccccctccag	gccatgagga	ttctgcagtt	aatcctgctt	gctctggcaa	4261
	caggtagcga	ggggatgggg	gcagggcagg	atcctccctc	ttgaatctct	gggatccctt	4321
50	aacctctgtg	gtctggacag	tgacagggct	gattccaaat	tacagaacaa	cccataaggc	4381
	acctgaactg	gagcagtggt	catgagggcc	ttgatgccct	tctagataat	ccctttaa	4441
	gccaaggag	gagaggtcaa	gggggtcgta	aagggtcccg	tggaggggct	gaggaagctg	4501
	gagttggggg	agcagtcact	caaagcgccc	aggacagggg	ctactgacca	accagtatgg	4561
	aagtatttcc	tttttttttt	ttcccagaga	caaagtcttg	ctctattgtc	caggctggag	4621
55	tgccgtggtg	ccaacacggc	tcactgcagt	cttgacttcc	cgggcttaag	tgatccttaa	4681
	gccatctcag	cttcccgggt	agctgggacc	acaggcacct	gccaccaagc	caggctaatt	4741
	gtttaattgt	ttgtagagat	gggggaggag	gtctcactat	gtttgcctag	gctgatctag	4801
	aactcctggg	ctcaagtaat	cctccacact	tagcctctca	aagtgtggg	attacaggca	4861
	tgagccactg	catttgacct	tatggaagta	ttttcatcct	ttaatacccg	acccagcat	4921



	ccagggcaac	ccagagggac	accagaccag	ggcccagacc	accactctc	tttctctcct	4981
	ccccacccc	atttctggga	gtcctcctgg	tctaccacct	ctccttctcg	agcccttct	5041
	tttgctctca	ccccctccag	ggcttgtagg	gggagagacc	aggatcatca	aggggttcga	5101
	gtgcaagcct	cactcccagc	cctggcaggc	agccctgttc	gagaagacgc	ggctactctg	5161
5	tggggcgacg	ctcatcgccc	ccagatggct	cctgacagca	gcccactgcc	tcaagccgtg	5221
	gggtgcgggg	ctggggcggt	gcgggggtgg	ggggctggga	atggggagat	ggatggagag	5281
	aagctcaggg	ataggggtgc	tggttaaggg	attagagatg	gggatgggta	gtgtcagcaa	5341
	ggttgatggg	ctcgagttgg	tattgaaggt	gggggatga	atggggttgg	gatggggcta	5401
	tggctgggaa	gggggcttcg	tggggagacc	tggaagaggt	tggaagcaga	gcatgttttc	5461
10	ttcatcctca	aaggtgtcac	tcacctctcc	cacctatgtc	tccccgacc	tttctcctc	5521
	caactactgt	ctctcccacc	tcagccgcta	catagtccac	ctggggcagc	acaacctcca	5581
	gaaggaggag	ggctgtgagc	agaccgggac	agccactgag	tccttcccc	accccggtt	5641
	caacaacagc	ctcccaaca	aagaccaccg	caatgacatc	atgctgtgtg	agatggcact	5701
	gccagtctcc	atcacctggg	ctgtgcgacc	cctcacccctc	tcctcacgct	gtgtcactgc	5761
15	tggcaccagc	tgcctcattt	ccggctgggg	cagcacgtcc	agccccagt	gtaggagcac	5821
	cagaggggaa	cctggcaggg	gggtgtgagg	agggagtggg	caggattgtg	gaagggttca	5881
	gggcatcaga	gatcgcggtc	acagtgacga	tgtgggataa	gttgagagga	tgtgtggaaa	5941
	acgtcaggat	aggggggtgg	ggacaaaagt	tggggccttg	gagtcagacg	gacgggatat	6001
	gcaatcatal	atccataacc	tcctggttgt	aagaccttag	gcaagcagct	tcacctctct	6061
20	gaatcttgat	tttcttctct	ataaaatgag	aatgattata	cccacctgtc	aggattggat	6121
	tagagataat	gtatatcaag	caactgacat	aaatcattta	ttggatagca	ggctgggcac	6181
	cgtggctcac	gcctgtaatc	ccagcacttt	gggaggccga	ggtgggaaga	tcacctgagg	6241
	tcaggacttt	gataccagcc	tgGCCAACgt	ggtgaaatcc	catctctact	aaaaatgtga	6301
	aaattagttg	ggcgtgggtg	tgtgcgcctg	taatccagc	tactcgggag	ggttaggcag	6361
25	gagaatcgct	taaacttggg	agacggaggt	tgcatgtagc	caagatcacg	ccactgcact	6421
	ccagcctggg	caacagagca	agactctgtc	tcgaaaaaaa	aaaaaaaaaa	gctggatagc	6481
	attgctgttg	ctattgtttac	aagaagagag	gtgaattggc	tgctgtctaa	gacagggatt	6541
	ccccagggg	cgggatcaca	gcaagcactg	cattagggga	ggtggcaggg	ggctcattcc	6601
	cacagcccct	cacgctgttt	ccacagtacg	cctgcctcac	accttgcgat	gcgccaacat	6661
30	caccatcatt	gagcaccaga	agtgtgagaa	cgcctacccc	ggcaacatca	cagacaccat	6721
	gggtgtgtgc	agcgtgcagg	aagggggcaa	ggactcctgc	caggtcagtg	tggtctccaa	6781
	ccacagcccc	atccccatcc	ccagcttcaa	tgacatcttt	accgacatcc	acaatttcat	6841
	ccccaacctc	aacccgcgga	cccctgaac	tcccaatcca	tctcttcccc	tgttcccggt	6901
	tctgacctca	gcacaaactt	cagctccatc	ccggtttcca	caccatttcc	agctccaacc	6961
35	atccccaaac	tcgtttttga	gcctaaccct	atcctttatc	ccacccataa	tcccagcttt	7021
	atcgctaaac	ctatcacctt	tcccagtgcc	taccatcctt	gtctcggccc	cactcctaag	7081
	caccgtcccc	acctcctccc	tggctaacac	catgctcaac	gctttctctg	accgacattc	7141
	tctctccccg	tgcccagggt	gactccgggg	gcctctgggt	ctgtaaccag	tctcttcaag	7201
	gcattatctc	ctggggccag	gatccgtgtg	cgatcaccgg	aaagcctggg	gtctacacga	7261
40	aagtctgcaa	atatgtggac	tggatccagg	agacgatgaa	gaacaattag	actggaccca	7321
	cccaccacag	cccatacccc	tccatttcca	cttggtgttt	ggttcctgtt	cactctgtta	7381
	ataagaaacc	ctaagccaag	acctctacg	aacattcttt	gggcctcctg	gactacagga	7441
	gatgtgtcoa	cttaataaat	aacctggggg	tcgaaatcag	tgagacctgg	attcaaatcc	7501
	tgccttgaaa	tattgtgact	ctgggaatga	caacacctgg	tttgttctct	gttgatatcc	7561
45	cagccccaaa	gacagctcct	ggccatatat	caagggtttca	ataaatattt	gctaaatgag	7621
	tgaatctact	gagtgtttac	tatgtgtctg	acctgatcc	aatggctttt	attttatttt	7681
	attttttgac	agagtctcgc	tctgtcaccc	aggtggagt	acagtgggtc	tatctctgct	7741
	cactgcaacc	tccacctcct	gggttcaagc	aattctcctg	cctcagcctc	ctgaatagct	7801
	gggattacag	gtgcctacca	ccacatccgg	ctaatttttg	tatttttttg	tagagatggg	7861
50	gcttcaccat	gttgccacgg	ctggtctcga	actcctgacc	tcagatgac	tgccctcctt	7921
	ggcctcccaa	actcctggga	ttacagacgt	gagccaccgc	gcccggccgg	cttctattta	7981
	ttaattaaaa	gaaattaaat	taattaatct	atttaggaga	cagtcttgct	ctggtgcccc	8041
	ggctggagtg	cagtaacaat	cacagctcac	ggcaatctca	atttctctgg	gtcaagtgat	8101
	tgtctccctt	cagcctccag	agtagctggg	actacaggca	catgccacga	agcccagcta	8161
55	atttttgtag	ttttogtaga	gacagaggtc	tcagtatgtt	gccccggcta	gtctcaaatc	8221
	cctgggtcca	agcagtctgt	cctcctcagc	ctccaaaagt	ggtagagatta	caggcatgag	8281
	tcgctgtgcc	tggcctccaa	gcactttcaa	atgtatcaac	ttaatcctca	caaaacctg	8341
	tgaggtcggt	actgttttca	tacctatttt	atagttgaag	aaacagacac	agagaagcaa	8401
	agtcacttgc	tcacagtcac	gtggctagga	gagcaaggat	ctgaagcaag	gcgatctctt	8461

```
aattaccaag tgatgttcct ggagtaaggo tctgtttggt tcctttcctg taaaatgctg 8521
catgcaaaag tataacacag taagtaaaga agtcagttag cctgcacata ctaagaccta 8581
accaaaggag cttattgttt ttctccaact tccatgatag gtaattagat agtggagacc 8641
tctgctggcc aatatggtag ccactaaccg cagctggctc ttccaattaa aattacataa 8701
5 agccagaaat gtaactcctc tgtctcactt gttatatctc caaggctgga tagccacatg 8761
tgactggtgg tggctggatt agctagtgcata taaaaacat cactgcagaa agttcagctg 8821
agcagcactg agttagatgg cctctgaaga ggatgtccca cggagagaat ccagaactca 8881
ggatcctttt tttttttttt ctttgcgaca gagtcttgct ctgtcaccca ggctggagtg 8941
cagtggcgtg atctcggtc actgcaactt ctgcctccca ggttcaagca attctcctgc 9001
10 ctcagcctcc ctagtagctg ggactacagg cctgtgccaa catcccagc taatttttgt 9061
gtcctttttag tagagatggg gtttcactat gttggccagg ctggtctcga actcctgacc
```